



**Universidad Nacional Mayor de San Marcos**

**Universidad del Perú. Decana de América**

Dirección General de Estudios de Posgrado

Facultad de Ciencias Matemáticas

Unidad de Posgrado

**"Modelo autorregresivo con heterocedasticidad  
condicionada generalizada fraccionalmente integrado"  
Caso: Estimación de la volatilidad del tipo de cambio  
nominal del Perú**

**TESIS**

Para optar el Grado Académico de Magíster en Estadística

Matemática

**AUTOR**

José Luis BRIONES ZÚÑIGA

**ASESOR**

Mg. Antonio BRAVO QUIROZ

Lima, Perú

2018



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

## Referencia bibliográfica

---

Briones, J. (2018). *"Modelo autorregresivo con heterocedasticidad condicionada generalizada fraccionalmente integrado" Caso: Estimación de la volatilidad del tipo de cambio nominal del Perú.* [Tesis de maestría, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Matemáticas, Unidad de Posgrado]. Repositorio institucional Cybertesis UNMSM.

---

## ACTA DE SUSTENTACIÓN DE TESIS DE GRADO ACADÉMICO DE MAGÍSTER

Siendo las, 17:05 horas del día miércoles diecinueve de setiembre del dos mil dieciocho, en el Auditorio de la Facultad de Ciencias Matemáticas, el Jurado Evaluador de Tesis, Presidido por la Dra. Rosa María Inga Santivañez e integrado por los siguientes miembros, Mg. Wilfredo Eugenio Domínguez Cirilo (Jurado Informante), Mg. Olga Lidia Solano Dávila (Jurado Evaluador), Dra. Ilse Janine Villavicencio Ramírez (Jurado Evaluador) y el Mg. Antonio Bravo Quiroz como Miembro Asesor, se reunieron para la sustentación de la (tesis titulada: «MODELO AUTORREGRESIVO CON HETEROCEDASTICIDAD CONDICIONADA GENERALIZADA FRACCIONALMENTE INTEGRADO. CASO: ESTIMACIÓN DE LA VOLATILIDAD DEL TIPO DE CAMBIO NOMINAL DEL PERÚ» presentada por el Bachiller José Luis Briones Zúñiga, para optar el Grado Académico de Magíster en Estadística Matemática.

Luego de la exposición del graduando, los Miembros del Jurado hicieron las preguntas correspondientes, así como las observaciones e inquietudes acerca del trabajo de tesis, a las cuales el Bachiller José Luis Briones Zúñiga respondió con acierto y solvencia, demostrando pleno conocimiento del tema.

A continuación se realizó la calificación correspondiente, según tabla adjunta, resultando el Bachiller José Luis Briones Zúñiga aprobado con el calificativo de Bueno.....  
DIECISEIS.....(16).....

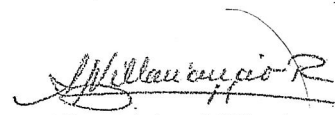
Habiendo sido aprobada la sustentación de la Tesis, el Jurado Evaluador recomienda para que el Consejo de Facultad apruebe el otorgamiento del grado académico de Magíster en Estadística Matemática al Bachiller José Luis Briones Zúñiga.

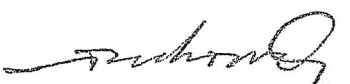
Siendo las 18:10 horas, se levantó la sesión, firmando para constancia la presente Acta.

  
Dra. Rosa María Inga Santivañez  
Presidenta

  
Mg. Wilfredo Eugenio Domínguez Cirilo  
Miembro

  
Mg. Olga Lidia Solano Dávila  
Miembro

  
Dra. Ilse Janine Villavicencio Ramírez  
Miembro

  
Mg. Antonio Bravo Quiroz  
Miembro Asesor

## **FICHA CATALOGRÁFICA**

**JOSÉ LUIS BRIONES ZÚÑIGA**

**"MODELO AUTORREGRESIVO CON HETEROCEDASTICIDAD  
CONDICIONADA GENERALIZADA FRACCIONALMENTE  
INTEGRADO"**

**CASO: ESTIMACIÓN DE LA VOLATIDAD DEL  
TIPO DE CAMBIO NOMINAL DEL PERÚ, (Lima) 2018**

**X.,133p., 29.7 cm (UNMSM, Magister, Estadística Matemática, 2018)  
Tesis de Maestría, Universidad Nacional Mayor de San Marcos.  
Facultad de Ciencias Matemáticas.  
Unidad de Posgrado, UNMSM/FCM.**

**”Modelo Autorregresivo con Heterocedasticidad Condicionada  
Generalizada Fraccionalmente Integrado”**

**Caso: Estimación de la volatilidad del tipo de cambio nominal del  
Perú**

**José Luis Briones Zúñiga**

Tesis presentada a consideración del jurado examinador nombrada por la Unidad de Posgrado de la Facultad de Ciencias Matemáticas de la Universidad Nacional Mayor de San Marcos, como parte de los requisitos para optar el Grado Académico de Magíster en Estadística Matemática.

Aprobada por

Dra. Rosa María Inga Santivañez  
**Presidenta**

Mg. Wilfredo Eugenio Dominguez Cirilo  
**Miembro**

Mg. Olga Lidia Solano Dávila  
**Miembro**

Dra. Ilse Janine Villavicencio Ramírez  
**Miembro**

Mg. Antonio Bravo Quiroz  
**Miembro Asesor**

*A Dios con el todo es posible, al Ing. Benavides padrino de mi madre, a mis  
padres y a mi asesor Mg. Antonio Bravo Quiroz.*

*Un ser humano es parte del todo que llamamos Universo,  
una parte limitada en el tiempo y el espacio.*

*Está convencido de que él mismo, sus pensamientos  
y sentimientos son algo independiente de los demás.*

*Esto es una ilusión óptica de la conciencia, una cárcel  
para nosotros que nos limita a nuestros deseos  
personales y a sentir afecto solo por los más cercanos.*

*Nuestra tarea es liberarnos de esa cárcel, ampliando  
nuestro círculo de compasión, hasta abarcar todos los  
seres vivos y toda la naturaleza.*

*Albert Einstein..*



# Índice general

|  |           |
|--|-----------|
| <b>1. Introducción</b>   | <b>1</b>  |
| 1.1. Situación Problemática . . . . .  | 4         |
| 1.2. Formulación del Problema . . . . .  | 6         |
| 1.3. Objetivos . . . . .   | 7         |
| 1.3.1. Objetivo General . . . . .  | 7         |
| 1.3.2. Objetivos Específicos . . . . .   | 7         |
| <b>2. Marco Teórico</b>  | <b>8</b>  |
| 2.1. Teoría del Caos . . . . .   | 8         |
| 2.2. Enfoque reducción Probabilística . . . . .  | 10        |
| 2.3. Nociones Básicas de procesos estocásticos . . . . .   | 12        |
| 2.3.1. Introducción . . . . .  | 12        |
| 2.3.2. Propiedades de los procesos Gaussianos . . . . .  | 14        |
| 2.3.3. Propiedades generales de distribuciones de procesos . . . . .   | 17        |
| 2.3.4. Movimiento Browniano . . . . .  | 19        |
| 2.3.5. Movimiento Browniano Fraccional . . . . .   | 27        |
| 2.3.6. Volatilidad Dinámica . . . . .  | 45        |
| 2.4. Análisis Estadístico para series de tiempo con memoria larga . . . . .                                      | 62        |
| 2.4.1. Análisis Fractal . . . . .  | 62        |
| 2.4.2. Análisis Estadístico . . . . .  | 62        |
| 2.4.3. Pruebas de Memoria Larga . . . . .  | 71        |
| 2.4.4. Métodos Semi-Paramétricos . . . . .   | 76        |
| 2.4.5. Estimación del modelo GARCH fraccionalmente integrado (FIGARCH) . . . . .                                 | 77        |
| <b>3. Metodología</b>  | <b>80</b> |
| 3.1. Diseño Metodológico . . . . .   | 80        |
| 3.1.1. Tipo de investigación . . . . .   | 80        |
| 3.1.2. Metodología Econométrica . . . . .  | 81        |
| 3.1.3. Modelo autoregresivo con heterocedasticidad condicionada generalizada fraccionalmente integrado . . . . . | 81        |
| 3.1.4. Ruido Blanco Fraccionario . . . . .   | 89        |
| 3.1.5. La no negatividad de la varianza condicionada . . . . .   | 90        |
| 3.1.6. La función Impulso respuesta . . . . .  | 92        |
| 3.1.7. Estimación y Pronóstico . . . . .   | 97        |
| 3.2. Población y tamaño de muestra . . . . .   | 100       |
| 3.2.1. Población . . . . .   | 100       |
| 3.2.2. Muestra . . . . .   | 100       |

|   |            |
|---|------------|
| <b>4. Resultados</b>  | <b>101</b> |
| 4.1. Técnicas de recolección de datos . . . . .                 | 101        |
| 4.2. Técnicas para el procesamiento de la Información . . . . . | 101        |
| 4.2.1. Análisis Estadístico . . . . .                           | 101        |
| 4.2.2. Pruebas de Memoria Larga . . . . .                       | 112        |
| 4.3. Resultados Metodológicos . . . . .                         | 118        |
| <b>5. Conclusiones</b>  | <b>124</b> |
| 5.1. Conclusiones . . . . .                                     | 124        |
| 5.2. Sugerencias para futuras investigaciones . . . . .         | 126        |
| 5.3. Apéndice 1: BDS . . . . .                                  | 130        |
| 5.3.1. Prueba de no linealidad (BDS) . . . . .                  | 130        |
| 5.4. Apéndice 2 . . . . .                                       | 132        |
| 5.4.1. Prueba de Ljung Box Portmanteau . . . . .                | 132        |
| 5.5. Apéndice 3 . . . . .                                       | 132        |
| 5.5.1. Prueba de normalidad de Jarque Bera . . . . .            | 132        |
| 5.6. Apéndice 4 . . . . .                                       | 133        |
| 5.6.1. Criterios de información . . . . .                       | 133        |

# Índice de figuras

|  |     |
|--|-----|
| 2.1. Atractor Extraño . . . . .  | 9   |
| 2.2. Función de densidad de probabilidad para $n=3$ . . . . .                                      | 64  |
| 2.3. Función de densidad de probabilidad para $n=10$ . . . . .                                     | 64  |
| 2.4. Función de densidad de probabilidad para $n=100$ . . . . .                                    | 65  |
| 2.5. Función de densidad de probabilidad para $n=500$ . . . . .                                    | 65  |
| 2.6. Distribución Normal . . . . .   | 68  |
| 2.7. Autosimilitud a Escala . . . . .  | 70  |
| 2.8. Ruido Blanco (Movimiento Browniano) . . . . .   | 73  |
| 2.9. Ruido Negro (Persistencia) . . . . .  | 74  |
| 2.10. Ruido Rosa (Antipersistencia) . . . . .  | 74  |
| 4.1. Serie tipo de cambio . . . . .  | 102 |
| 4.2. Retorno porcentual de la serie . . . . .  | 102 |
| 4.3. Cuadrado de los retornos porcentuales de la serie . . . . .                                   | 103 |
| 4.4. Valor absoluto de los retornos de la serie . . . . .  | 103 |
| 4.5. Principales Estadísticos . . . . .  | 107 |
| 4.6. Histograma de Densidad de Kernel para retornos porcentuales . .                               | 108 |
| 4.7. Histograma de Densidad de Kernel para retornos porcentuales al<br>cuadrado . . . . .          | 108 |
| 4.8. Correlogramas para los retornos porcentuales . . . . .  | 109 |
| 4.9. Correlogramas para los retornos porcentuales al cuadrado . . . . .                            | 110 |
| 4.10. Correlogramas para los retornos porcentuales en valor absoluto . .                           | 110 |
| 4.11. Ajuste del logaritmo de reescalamiento de la serie del cuadrado de<br>los retornos . . . . . | 116 |
| 4.12. Ajuste del logaritmo de reescalamiento de la serie $r$ . . . . .                             | 117 |
| 4.13. Retorno porcentuales del tipo de cambio . . . . .  | 118 |
| 4.14. Funcion de autocorrelación del cuadrado de los retornos . . . . .                            | 119 |
| 4.15. La varianza condicional estimada . . . . .   | 122 |

# Índice de cuadros

|  |     |
|--|-----|
| 4.1. Resumen descriptivo de los retornos porcentuales . . . . .  | 104 |
| 4.2. Autocorrelacion de los retornos porcentuales . . . . .  | 105 |
| 4.3. Autocorrelación de los retornos porcentuales al cuadrado . . . . .                                    | 106 |
| 4.4. Prueba de BDS de retornos porcentuales para distintas dimensio-<br>nes . . . . .                      | 107 |
| 4.5. Pruebas para retornos porcentuales . . . . .  | 112 |
| 4.6. Pruebas para retornos porcentuales al cuadrado . . . . .  | 112 |
| 4.7. Pruebas para retornos porcentuales en valor absoluto . . . . .  | 112 |
| 4.8. Pruebas para retornos porcentuales . . . . .  | 113 |
| 4.9. Pruebas para retornos porcentuales al cuadrado . . . . .  | 114 |
| 4.10. Cifras de rango reescalado del cuadrado de los retornos . . . . .                                    | 115 |
| 4.11. Cifras de rango reescalado de $r$ . . . . .  | 117 |
| 4.12. Coeficiente de Hurts . . . . .   | 117 |
| 4.13. Estimación de modelos FIGARCH . . . . .  | 120 |
| 4.14. Criterio de información para la elección del modelo . . . . .  | 120 |
| 4.15. Medidas de error de los pronósticos de la varianza condicional . .                                   | 121 |
| 4.16. Q-Statistics on Squared Standardized Residuals Ljung-Box . . . .                                     | 121 |
| 4.17. ARCH test on Squared Standardized Residuals . . . . .  | 121 |
| 4.18. Correlograma para Std.Residuales al cuadrado utilizando desvia-<br>ción típica $1/T^{0,5}$ . . . . . | 123 |

# Resumen

En la presente tesis se analiza el cambio de paradigma de un movimiento browniano ordinario a un movimiento browniano fraccional en el proceso de la volatilidad del tipo de cambio, es decir el elemento de persistencia en una serie caótica muy sensible a cambios en las condiciones iniciales, los cuales generan impactos decisivos en la dinámica de su movimiento de esta manera identificándose patrones en su conducta a primera vista aleatoria pero fractalmente con un patrón a modelar, se demuestra que la serie de tiempo sujeto de estudio es un proceso con incrementos no estacionarios y dependientes distinguidas por la no linealidad negando la posibilidad de ser un proceso martingala, debido a la evidencia del coeficiente de Hurts y otras pruebas semiparamétricas que la respaldan por lo que se demuestra también que la variación cuadrática del proceso es cero. Por otro lado se muestra que dicha persistencia tiende a desaparecer de manera hiperbólica para ello se utilizó la función impulso respuesta acumulativa también llamada memoria larga. Por lo tanto el centro neurálgico de esta investigación es la persistencia en modelos no lineales heterocedásticos (FIGARCH) Modelos autoregresivos con heterocedasticidad condicionada generalizada fraccionalmente integrados. Utilizando las principales propiedades de procesos gaussianos y los casos específicos de movimiento Browniano y movimiento Browniano Fraccional.

Para dicha aplicación se utilizó a la variable tipo de cambio y mediante modelos de series de tiempo de memoria larga poder analizar la persistencia del efecto existente en la volatilidad de dicha serie, Considerando que existen muchos puntos de vistas acerca del estudio de este fenómeno caótico, en la presente tesis se logra la formulación y estimación econométrica para entender de mejor manera la naturaleza de un indicador macroeconómico clave en la toma de decisiones estatales. De esta manera poder demostrar la dependencia de los intervalos ajenos si importar su distancia en un proceso estocástico, es decir la existencia de la dependencia no lineal entre los incrementos de una serie de tiempo.

Palabras Claves:

Memoria Larga, FIGARCH, Movimiento Browniano Fraccional, Función Impulso respuesta , ARCH-GARCH, Movimiento Browniano.

# Abstract

In the present thesis the change of paradigm of a movement is analyzed ordinary Brownian to a fractional Brownian movement in the process of volatility of the exchange rate, ie the element of persistence in a series chaotic, very sensitive to changes in the initial conditions, which generate decisive impacts on the dynamics of their movement in this way identifying patterns in their behavior at first glance randomly but fractally with a pattern to model, it is shown that the time series subject of study is a process with non-stationary and dependent increases distinguished by non-linearity denying the possibility of being a martingale process, due to the evidence of Hurts coefficient and other semi-parametric tests that support it by what is also shown that the quadratic variation of process is zero. On the other hand it shows that such persistence tends to disappear in hyperbolic way for it was used or the impulse response function cumulative also called long memory. Therefore the neuralgic center of this research is the persistence in nonlinear heteroscedastic models (FIGARCH) Autoregressive models with conditioned heterocedasticity fractionally integrated generalized. Using the main properties of Gaussian processes and the specific cases of Brownian movement and Fractional Brownian movement.

For this application, the variable exchange rate and through Time series models of long memory can analyze the persistence of the existing effect on the volatility of said series, considering that there are many points of view about the study of this chaotic phenomenon, in the This thesis is achieved the formulation and econometric estimation to understand better the nature of a key macroeconomic indicator in the State decision making. In this way, we can demonstrate dependence of the intervals of others if importing their distance in a stochastic process, is say the existence of non-linear dependence between the increments of a Time series.

Keywords: Long Memory, FIGARCH, Fractional Brownian Movement, Function Impulse response, ARCH-GARCH, Brownian Movement.

# Capítulo 1

## Introducción

El tipo de cambio es una variable macroeconómica cuya dinámica depende tanto de factores externos como internos como también de las políticas económicas adoptadas y la sensibilidad ante decisiones exógenas que afectan en cierta medida a la economía de un país. En el caso peruano de manera intuitiva se denota la complejidad del sistema en la cual actuamos y el caos inherente en el mismo. Esta investigación tendrá por objetivo estudiar estas variaciones por medio de los clásicos modelos de series de tiempo que puedan captar la memoria de corto plazo muy utilizados en la actualidad y de largo plazo motivo de la tesis, señalando en el primer Grupo modelos ARCH y GARCH y en el segundo como FIGARCH centrándonos en la no linealidad de FIGARCH, es decir dar paso a un sistema económico caótico con reglas distintas a tratar.

Con el desarrollo de la tecnología y la inclusión económica (Globalización), nuestro sistema económico ha desarrollado una clara sensibilidad ante factores sean interno y externos, ya sean de gran o menor magnitud, desarrollando un entorno caótico y multivariable, incrementando de esta manera la entropía intrínseca de la realidad, por tal motivo se requieren de técnicas Estadísticas y de modelos multifractales que reflejen la dinámica del sistema económico de nuestro país, con el objetivo de analizar dichos fenómenos sirviendo de instrumento para una adecuada planificación. Por tal motivo se necesita un análisis profundo de la naturaleza de dichas variables para detectar patrones, comportamientos, tendencias, que nos permitan tomar decisiones en relación a la perspectiva de la dinámica de la economía del Perú. Estos factores adquieren cierto grado de incertidumbre o riesgo que por lo general es mayor en países del tercer mundo debido a la falta de institucionalidad y una adecuada estructura ética, ocasionando inestables escenarios económicos, políticos y sociales.

De lo anterior, se desprende el hecho de porque estudiar modelos de especificación adecuada en volatilidad del tipo de cambio, para lo cual en este trabajo se abordara de la siguiente manera:

Como primera etapa se desarrollara los conceptos básicos de procesos estocásticos, teoremas importantes y necesarios. Para poder entender la naturaleza del sistema objeto de estudio, es importante mencionar que la definición en este

sección de procesos estocásticos es basada en la dada por Paul Levy (1955) para luego desarrollar especificaciones de procesos gaussianos como el movimiento browniano, movimiento browniano fraccional extensiones propias del browniano estándar los cuales son esenciales para diferenciar entre los modelos de memoria corta y memoria larga, de esta manera sentar las bases teóricas y poder desarrollar el modelo de FIGARCH y así capturar y cuantificar las dependencias temporales observadas en las dinámicas de la economía peruana. En particular poder analizar e interpretar las características del modelo FIGARCH el cual permite observar meticulosamente la caída hiperbólica del cuadrado de los rezagos absolutos de las innovaciones de la función de la varianza condicional.

La siguiente etapa tiene por objetivo abordar los modelos ARCH de Robert F. Engle en 1982 y GARCH de Bollerslev (1991) simétricos (reacción de la volatilidad ante innovaciones positivas y negativas) mostrando los conceptos teóricos básicos, métodos de estimación de los parámetros mediante el algoritmo de Scoring Fisher los cuales generaron un importante impacto en el estudio de los modelos de volatilidad ampliando al campo de las Estructuras cuadráticas basados en gran parte en los modelos ARIMA. Dichos métodos Heterocedásticos surgieron como necesidad de modelar las oscilaciones a corto plazo que recoge la varianza condicionada de las series de tiempo que la Metodología Box-Jenkins consideraban estacionarias y varianza constante los cuales son usados para obtener un pronóstico óptimo de valores futuros de la serie.

Por lo cual es que a la par de la evolución de la metodología de series de tiempo han surgido problemas que los economistas tratan de explicar, uno de estos es porque muchas variables de corte financiero y en nuestro caso el tipo de cambio (sol - dólar), son difíciles de modelar y por ende de pronosticar, por su posible sensibilidad a condiciones iniciales. Esta condición medida por diferentes técnicas surgidas de la teoría del caos y en particular al hidrólogo Hurst H.E., que con su trabajo de 1951 revolucionó la forma de estudiar los modelos serie de tiempo mediante el estudio de procesos con una caída hiperbólica de la correlación de los intervalos ajenos son importar el tiempo (memoria de largo plazo) esto es persistencia, supuso que las condiciones iniciales de la serie afecta no solo al dato posterior inmediato Markov, sino que la afección inicial es persistente durante toda la serie, es decir son series caóticas, más tarde esta idea fue retomada por Hosking en 1981, Granger y Joyeux en el año de 1980, dado así a conocer la versión discreta del Movimiento Browniano Fraccional y que actualmente se conoce como modelos ARFIMA o FARIMA, la característica principal es que la serie es transformada de acuerdo al valor de "d", el cual es un número contenido en los números reales.

De la idea anterior surgió también el modelo FIGARCH, propuesto inicialmente por Bollerslev, Baillie y Mikkelsen en 1996, Granger y Ding en 1996 y Chung en 1999, se basa en el hecho de pensar que los retornos cuadráticos poseen memoria larga y por lo tanto que estos impactan en la varianza condicional, siendo esta persistente y mucho mayor que el modelo ARFIMA, por lo que el impacto de un shock desaparece muy lentamente manifestándose en la lenta convergencia de la estabilización de la varianza condicionada.



Para probar el fenómeno de memoria larga en los modelos serie de tiempo y como consecuencia obtener el orden de integración fraccional de la misma se recurrirá a pruebas básicas como son el coeficiente de Hurst y otras pruebas semi paramétricas como el estimador local de Whittle y la prueba Geweke y Porter-Hudak o GPH.

Como siguiente etapa, tendrá por objetivo realizar un análisis de los modelos propuestos anteriormente y se dará las conclusiones de los mismos, generando sugerencias de los mejores modelos que permiten ver la dinámica del shock su velocidad de convergencia a través de los coeficientes de la función impulso respuesta acumulativa.

Por lo tanto la estructura de la tesis viene dada por los siguientes capítulos: En el capítulo 2, se presentan los conceptos teóricos básicos más importantes para el posterior desarrollo del modelo FIGARCH, conceptos y características de un sistema caótico, la formulación de un modelo econométrico desde el enfoque de reducción probabilística, nociones básicas de procesos estocásticos, propiedades de procesos gaussianos y ejemplos particulares de procesos gaussianos como son movimiento browniano estándar y movimiento browniano fraccional. Para luego definir la variable a modelar la volatilidad sus propiedades y antecedentes de la modelación FIGARCH, como son ARCH Y GARCH.

En el capítulo 3 se empieza a desarrollar la metodología econométrica FIGARCH partiendo de ARCH pasando por GARCH y llegando a FIGARCH, describiendo algunas de sus propiedades más importantes que me ayudaran a cuantificar y analizar la persistencia de la durabilidad del shock del tipo de cambio lo cual nos ayudará a comprender el grado de sensibilidad de la estabilidad paramétrica en la memoria larga del modelo construido. Por otro lado también se realizará la aplicación de la técnica econométrica a la serie de tipo de cambio de los ultimos 20 años.

En el capítulo 4, se mostrarán los resultados de la estimación del modelo de memoria larga, se realizará la interpretación de los mismos.

En el capítulo 5, se mostrarán las conclusiones y recomendaciones del modelo no lineal.

## 1.1. Situación Problemática

En el campo de las Ciencias Sociales específicamente en la Economía es notorio la ausencia de teorías que expliquen de manera satisfactoria la dinámica de sus fenómenos de estudio en comparación con las Ciencias Naturales. *A mediados de los 60, una queja que oía con frecuencia entre los economistas era: "Dado que sabemos que tal y tal expresión estadística muestra un comportamiento convergente en todos los otros campos de la ciencia, ¿Cómo puede ser que mi propio campo (se quejaba mi interlocutor) afronte el sólo la maldición de tener que enfrentar expresiones estadísticas divergentes? Cuando los otros campos de la ciencia pueden ser estudiados con métodos matemáticos bien probados que vienen en manuales familiares ¿Porqué tiene que pasar que en mi propio campo necesitemos técnicas nuevas, para las que las únicas referencias son siempre tomos polvorientos escritos en francés, o incluso en polaco, o incomprensibles monografías modernas?"*. (Mandelbrot, 1985:215).

Un argumento muy utilizado es la diferencia de edad, pero si tomamos en cuenta que la teoría de probabilidades surgió asociada al análisis de fenómenos de las ciencias Sociales más de un siglo antes que las investigaciones y teorías analíticas de la calor de Fourier y la energética de Helmholtz pasando por la mecánica estadística de Maxwell, Boltzman y Gibbs para luego llegar a la mecánica cuántica de Planck, Einstein, Bohr, Schrodinger y Heisenberg. Por lo tanto no sería un argumento sólido la distancia temporal entre la economía y la física indeterminista. Entonces surge en este contexto la figura del matemático polonés de ascendencia judía Benoit Mandelbrot resultado del vasto conocimiento de la escuela parisina de Emilie Borel, Henri Lebesgue, Louis Bachelier, Maurice Frechet, Paul Lévy, entre otros, a su vez herederos de la figura de Henri Poincare el cual junto con Lyapunov se les consideran los precursores de la matemática indeterminista entre 1889 y 1900. En el año de 1961, Mandelbrot como resultado de sus investigaciones económicas sobre los modelos estadísticos de la distribución de los tamaños del ingreso personal en una economía de mercado capitalista (Basados en los trabajos de Wilfredo Pareto, 1896), planteó la causa de tal diferencia a la predominancia de las distribuciones de Pareto en los fenómenos básicos de las ciencias sociales denominado por Mandelbrot como distribuciones Lévy-Pareto, distribuciones estables de Lévy conocidos por los matemáticos y distribuciones de Pareto según la ciencia económica. Alejándose de esta manera del caso gaussiano.

Las distribuciones de Lévy-Pareto se caracterizan por ser jerarquizada no ergódica es decir las variables que construyen al proceso estocástico tienen diferentes pesos específicos en el resultado final de la suma. Por lo tanto los efectos del impacto de ciertas realizaciones se diluyen muy lentamente de modo que la sumatoria de las variables aleatorias no causen la anulación rápida de este impacto como lo exige el teorema del límite central.

De este modo Mandelbrot diferencia dos clases de indeterminismo (Azar), el indeterminismo primario o Benigno y el indeterminismo de segunda especie o azar salvaje (Azar puro). El primero asociado a las obras de Laplace y Gauss, *"Todo el valor epistemológico de la teoría de probabilidades se basa en esto: los fenómenos*

*aleatorios considerados en su acción colectiva a gran escala, crean una regularidad no aleatorio”.* (B.V. Gnedenko y A.N. Kolgomorov, *limit distributions for sums of independent random variables*). Dichas fluctuaciones de orden primario reposan sobre los pilares de la ley de los grandes números o denominado también teorema ergódico y el teorema del límite central generando de esta forma el ruido blanco gaussiano que admite la existencia de memoria como un fenómeno mínimo de corto plazo o correlaciones seriales a mínima escala que tanto éxito alcanzaron en el estudio de fenómenos de dinámica determinístico no lineal a partir de la mitad del siglo XX.

El indeterminismo de segundo orden o ruido fraccionario contradice a la teoría econométrica ortodoxa en la que la memoria a largo plazo toma la forma de intensas fluctuaciones de baja frecuencias correlacionadas entre sí para cualquier escala temporal.

En este sentido *”Por lo tanto en líneas generales, un patrón de sucesos es científicamente significativo y puede ser considerado con posibilidades de tener continuidad, sólo si en cierto sentido su probabilidad de haber ocurrido por casualidad es muy pequeña...Pero cuando se trabaja en un campo en el que el ruido de fondo es paretiano, debemos reconocer que la carga de pruebas con la que nos enfrentamos se encuentra más cerca de aquélla característica de la historia y la autobiografía que de la física...Algunos modelos estocásticos... son capaces de generar trayectorias en las que tanto el ojo entrenado como el profano distinguen el tipo de detalle que generalmente asociamos con las relaciones causales... pero estas estructuras deben ser consideradas ilusiones perceptivas”.* (Mandelbrot, 1963b:433-34).

## 1.2. Formulación del Problema

La modelización del proceso de volatilidad del tipo del cambio es una pieza que da solidez a la estructura de política económica de una nación generando bases para una dinámica democrática necesarias para el desarrollo de un sistema pero debido a la naturaleza heterocedástica de los retornos de la tasa de cambio de corto plazo, los clásicos modelos que suponen estacionariedad y homocedasticidad no son los más idóneos para modelarla. Siendo dicha modelización de gran importancia en la toma de decisiones macroeconómicas y en las últimas décadas motivo de investigaciones en el mercado financiero. En un inicio se consideraban los modelos clásicos de series e tiempo subyugados al supuesto de varianza constante o quizá poca variabilidad cuando en realidad su naturaleza es heterocedástica y caótica dejando entonces atrás el enfoque determinista por tal motivo es importante la especificación de un modelo que capture la naturaleza de la volatilidad. Para modelar la dinámica de dicho proceso, se han propuesto diversos métodos estadísticos, dentro de los cuales la metodología ARCH ha jugado un papel importante.

En investigaciones como el de Baillie y Bollerslev (1989,1991), Hsieh(1988), Engle(1990), Vlaar y Palm(1993), entre otros, han utilizado la metodología ARCH y GARCH para la modelización del proceso de volatilidad con la estructura econométrica del modelo gaussiano GARCH(1,1) suponiendo a la innovación unitaria de la varianza condicionada como una distribución gaussiana mesocurtica y simétrica, por tal motivo dicho modelo es el más empleado para estudiar la volatilidad, sin embargo un número importante de estudios, empleando base de datos diarias han encontrado evidencia sobre la persistencia en la volatilidad es decir dicha innovación es más parecida a una distribución T o una Levy Pareto con colas más pesadas en relación a una distribución normal aproximando el proceso por medio de modelos IGARCH, FIGARCH e HYGARCH.

Con especial atención en FIGARCH se caracteriza por la no estacionariedad del proceso de volatilidad , Davidson(2004), debido a que es un proceso autosimilar fractal. *En este sentido como resultado del incumplimiento de los supuestos de normalidad y linealidad encontrados en los sistemas dinámicos siendo estos supuestos simplificadores para los cuales se tienen soluciones estables de esta manera manteniendo a la ciencia económica en la seguridad del determinismo evitando la ruptura con el pensamiento causal de Laplace(1749-1827),(R.S.Garcia,2014),* esta tesis busca principalmente investigar y analizar al modelo FIGARCH como una posibilidad de explicar gran parte de la diversidad fenomenológica de las series económicas y financieras por tanto también especificar y cuantificar el proceso de volatilidad del tipo de cambio del sol peruano con respecto al dólar americano para el periodo comprendido entre Enero del 1997 hasta Marzo del 2017.

Por tal motivo se busca el modelo que mejor explica el comportamiento de los retornos de la tasa de cambio siendo indispensable para obtener el pronóstico de la volatilidad asociada a los mismos.

## **1.3. Objetivos**

### **1.3.1. Objetivo General**

El presente estudio, tiene como objetivo general, determinar la durabilidad de los shocks en la volatilidad de la serie del tipo de cambio en Perú del periodo (1997-2017), esto quiere decir, estimar el mejor modelo para la data existente y contraste acerca de sus parámetros y comportamiento, de tal forma de encontrar evidencia de la persistencia en el largo plazo de dichos impactos.

### **1.3.2. Objetivos Específicos**

1. Identificación el movimiento browniano fraccional en series de tiempo que se consideraban movimiento browniano ordinario.
2. Estimar y Especificar el modelo Autoregresivo con Heterocedasticidad Generalizada Fraccionalmente Integrado (FIGARCH) aplicado al tipo de cambio Sol- Dólar.
3. Cuantificar la persistencia del proceso de volatilidad de la serie a través de la función Impulso respuesta Acumulativa del modelo FIGARCH.
4. Desarrollar el modelo Autoregresivo con Heterocedasticidad condicionada Generalizada Fraccionalmente Integrado (FIGARCH)
5. Demostrar la idoneidad del modelo FIGARCH en relación a otros modelos de heterocedasticidad condicionada para modelar la volatilidad del tipo de cambio a largo plazo.

# Capítulo 2

## Marco Teórico

### 2.1. Teoría del Caos

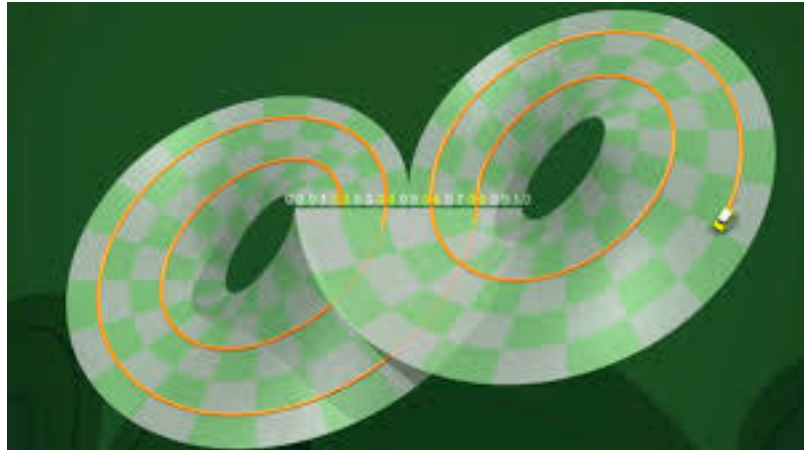
El precursor de la teoría del caos, fue el meteorólogo Edward Lorente (1917-2008). Estudiando la convección de la atmósfera mediante un sistema ecuaciones, se dio cuenta que con alteraciones mínimas de las variables iniciales, provocaban resultados muy divergentes. Este fenómeno se conoce como *Efecto Mariposa*, el cuál fue definido conceptualmente en un conferencia en el MIT a modo de pregunta *¿El aleteo de una mariposa en Brasil hace aparecer un tornado en Texas?*, Edward Lorente (1917-2008).

Una característica de las series caóticas es la dependencia sensible a condiciones iniciales, específicamente a mínimas condiciones iniciales generan grandes cambios en la dinámica a largo plazo del sistema este fenómeno es llamado Efecto Mariposa. Lorenz empezó a desarrollar su teoría del caos, y observó que aparentemente los sistemas de ecuaciones tienen un comportamiento aleatorio, pero mediante la geometría fractal se dibujaban una doble espiral repetidamente, denominada atractor de Lorenz muy parecido al símbolo de infinito. El atractor es uno de los conceptos fundamentales de la teoría del Caos, que se utiliza para representar la evolución en un sistema dinámico en un tiempo suficientemente largo. Geométricamente, un atractor puede ser un punto, una curva, una variedad en el caso de tratarse de un sistema estable, por lo contrario, si hablamos de un sistema caótico el atractor es una complicada estructura fractal conocido como atractor extraño.

En este caso de la series caóticas surgen atractores extraños, los cuales tienen una naturaleza fractal, es decir de fragmentación de irregularidad, tiene como característica que los puntos nunca se repiten y las orbitas nunca se interceptan. Por tal motivo se concluyen con tres características de los sistemas caóticos:

- Un excesiva sensibilidad a mínimos cambios en las condiciones iniciales generando grandes cambios en la dinámica a largo plazo del sistema.
- En términos de micro estado se refleja un movimiento browniano ordinario es decir una aleatoriedad, pero a nivel de macro estado se denota un cierto orden.

Figura 2.1: Atractor Extraño



- Son deterministas en tanto que existe alguna ecuación que gobierna su conducta.

Bart Kosko, autor de la llamada lógica borrosa, afirma de modo tajante que *cuanto más de cerca se mira un problema en el mundo real, tanto más borrosa se vuelve su solución.* Bart Kosko(1995).

Según la teoría de la totalidad un sistema es un fenómeno indisoluble único como un todo constituido por elementos interrelacionados caracterizados por la complejidad es decir concibe el mundo como un todo orgánico, fluido e interconectado

## 2.2. Enfoque reducción Probabilística

La Ciencia Económica como ciencia Social denota problemas epistemológicos y metodológicos diferentes a las ciencias naturales, por tal motivo hasta los años sesenta se consideró que la evidencia empírica no comprobaba la teoría planteada sino la falsificaba, entonces no se podía seguir el tradicional proceso de que la teoría debería preceder a la investigación empírica. Para luego empezar a desarrollarse metodologías en las cuales se toman en cuenta elementos de ambas partes de la teoría y de los datos, siguiendo el principio de Hegel .....*Hegel pensaba que la evolución de la idea se produce a través de un proceso Dialéctico, es decir un concepto se enfrenta a su opuesto y como resultado de este conflicto, se alza un tercero, la síntesis, la síntesis se encuentra más cargada de verdad que los dos anteriores opuestos*, Hegel(1770,1831), Entonces siguiendo esta línea de pensamiento se incorporan al modelo elementos de la muestra, por lo tanto el modelo no dependerá solamente del cuerpo de teoría ni de los datos, será un híbrido entre estos dos elementos una síntesis, es decir se tendrá en cuenta el proceso que genera los datos que se observan y sus desviaciones temporales.

En el año 1995 el método Hendry propone mediante la modelación econométrica como una alternativa de generar esta denominada síntesis de Hegel en donde se une la teoría económica y la evidencia que surge de los datos por lo tanto se requiere de construir un marco cuantitativo formal en donde se interrelacionen la teoría y evidencia extraída de los datos.

Pero es importante distinguir entre ambos extremos, es decir entre modelos basados en la teoría a priori y calibrados desde la evidencia de los datos y los modelos cuya estructura están en los datos, pues los primeros tienen una excesiva dependencia de la teoría pues dependen de la credibilidad de la teoría y los segundos modelos tienen una dependencia muestral. Sanos(1995), propone aplicar una metodología tradicional de datos experimentales a datos no experimentales es decir datos extraídos de proceso en curso no controlado, construir un modelo econométrico considerando la naturaleza no experimental de los datos que se están tratando. Considerando entonces un puente entre la teoría y los datos no experimentales (observacionales) este puente es denominado reducción probabilística como una estructura estadística de naturaleza no experimental, entonces el objetivo es especificar modelos estadísticos adecuados que capturen las características probabilísticas de los datos y proveer un resumen confiable de la información probabilística de los datos, de esta manera se mejorará la firmeza de la evidencia empírica como autoridad final de la validez de las teorías.

Entonces la reducción probabilística distingue dos mecanismos entre el modelo teórico es decir la aproximación a la realidad (cuerpo de teoría), proceso generador de datos y el modelo estadístico como un conjunto de supuestos probabilísticos de las variables aleatorias observables.

El modelo de reducción probabilística se puede resumir en cuatro etapas:

- Especificación: elección del modelo estadístico apropiado.
- Pruebas de mala especificación: Prueba de los supuestos al modelo elegido.
- Re-especificación: modelo alternativo en caso falle el modelo anterior.



- Identificación: Relación entre el modelo teórico y el modelo estadístico , mediante la re-parametrización.

## 2.3. Nociones Básicas de procesos estocásticos

### 2.3.1. Introducción

Si consideramos a los procesos estocásticos definidos en el espacio de probabilidad  $(\Omega, \mathfrak{F}, \mathbb{P})$ , donde  $\Omega$  es el espacio muestral no vacío,  $\mathfrak{F}$  es un sigma de algebra sub conjunto del espacio muestral no vacío y  $\mathbb{P}$  la medida de probabilidad una medida en  $\mathfrak{F}$ , conjunto de tiempo  $T$  y espacio de estados  $\mathbb{R}$ .

*Definición 1* Consideramos a una familia de variables aleatorias  $[X_t]_{t \in T}$  tal que para cada  $t \in T$  como un proceso estocástico, con espacio de probabilidad  $(\Omega, \mathfrak{F}, \mathbb{P})$ , conjunto de tiempo  $T$  y espacio de estados  $\mathbb{R}$ , entonces:

$$X_t : (\Omega, \mathfrak{F}, \mathbb{P}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

Se define a una variable aleatoria como una función real con valor finito. Sabiendo que los parámetros  $t \in T$  y  $\omega \in \Omega$  afectan a los procesos estocásticos, entonces :

$$X_t(\omega) \begin{cases} (t \text{ fijo}) & ; \text{Decimos } X_t(\omega) \text{ una variable aleatoria en un } (\Omega, \mathfrak{F}, P) \\ (\omega \text{ fijo}) & ; \text{Decimos } X_t(\omega) \text{ es una } \omega - \text{trayectoria en } \mathbb{R} \\ (t \in T \text{ y } \omega \in \Omega) & ; \text{Decimos Criterio de Levy} \end{cases}$$

Si  $t$  es fijo, entonces la función  $\omega \rightarrow X_t(\omega)$  es un a variable aleatoria con espacio de probabilidad  $(\Omega, \mathfrak{F}, P)$ .

Si  $\omega$  es fijo, entonces la función  $t \rightarrow X_t(\omega)$  es una función que toma valores en  $\mathbb{R}$  y se denomina  $\omega$ - trayectoria del proceso.

Entonces para construir a  $\mathcal{B}^L$  como la álgebra producto de conjuntos de  $\mathbb{R}^L$ , consideremos un  $L \subset T$ , en otras palabras  $\otimes_{t \in L} \mathcal{B}_t(\mathbb{R})$ , donde  $\mathcal{B}_t(\mathbb{R}) = \mathcal{B}(\mathbb{R})$  para cada  $t \in L$ .

Si  $\exists A_t \in \mathcal{B}(\mathbb{R})$  es un  $\sigma$ -álgebra, Entonces  $\Pi_{t \in L} A_t \subset \otimes_{t \in L} \mathcal{B}_t(\mathbb{R})$  y si  $\otimes_{t \in L} \mathcal{B}_t(\mathbb{R})$  es la mínima  $\sigma$ -álgebra del conjunto  $\mathbb{R}^L$  la cual es denotada por  $\mathcal{B}^L$ , entonces  $\{t \in L : A_t \neq \mathbb{R}\}$ , es finito.

Ahora para  $S \subset L \subset T$ , se la medida  $\varrho_S^L$  una función con dominio y rango:

$$\varrho_S^L : (\mathbb{R}^L, \mathcal{B}^L) \rightarrow (\mathbb{R}^S, \mathcal{B}^S)$$

o una proyección por:  $\varrho_S^L((f_t)_{t \in L}) = (f_t)_{t \in S}$

Donde  $f_t \in \mathbb{R} \forall t \in T$ , Entonces tenemos la medida  $\varrho_t^L$  una función con dominio y rango:

$$\varrho_t^L : \mathbb{R}^L \rightarrow \mathbb{R}$$

con  $t \in L$ , entonces:  $\varrho_t^L$ , si  $L=T$ .

Entonces  $\varrho_t^T$  la denotaremos como  $\varrho_t$ .

**Definición 2** Sea el proceso estocástico  $X = \{X_t\}_{t \in T}$  y tenemos  $F \subset T$ , se denota la función  $\varphi_X^F : (\Omega, \mathfrak{S}, \mathbb{P}) \rightarrow (\mathbb{R}^F, \mathcal{B}^F)$  entonces:

$\varphi_X^F(\omega) = \{X_t(\omega)\}_{t \in F}$ , si tengo  $[\varphi_X^F]^{-1}(\omega) \in \mathfrak{S} \forall \omega \in \mathcal{B}^F$ , entonces es medible con respecto a  $\mathfrak{S}$  y para cada  $t \in F$  se tiene que  $\pi_t^F \circ \varphi_X^F = X_t$  es medible para cada  $t \in F$ .

Si

$$\begin{aligned}\pi_t^F : (\mathbb{R}^F, \mathcal{B}^F) &\rightarrow (\mathbb{R}^T, \mathcal{B}^T) \\ \varphi_X^F : (\Omega, \mathfrak{S}, \mathbb{P}) &\rightarrow (\mathbb{R}^F, \mathcal{B}^F)\end{aligned}$$

Entonces  $\pi_t^F \circ \varphi_X^F : (\Omega, \mathfrak{S}, \mathbb{P}) \rightarrow (\mathbb{R}^F, \mathcal{B}^F) \rightarrow (\mathbb{R}^T, \mathcal{B}^T)$

Por lo tanto

$$\begin{aligned}(\pi_t^F \circ \varphi_X^F)^{-1}(A) &\in \mathfrak{S} \forall A \in \mathcal{B}^T \\ (\varphi_X^F)^{-1}(\pi_t^F)^{-1}(A) &\subset (\varphi_X)^{-1}(\mathcal{B}^F) \subset \mathfrak{S} \text{ Si} \\ \varphi_X : (\Omega, \mathfrak{S}, \mathbb{P}) &\rightarrow (\mathbb{R}^T, \mathcal{B}^T)\end{aligned}$$

Entonces  $\varphi_X^{-1} : \mathcal{B}^T \rightarrow [0, 1]$

Entonces tengo la función  $\mathbb{P} \circ \varphi_X^{-1} : \mathcal{B}^T \rightarrow [0, 1]$

$$[\mathbb{P} \circ \varphi_X^{-1}](A) = \mathbb{P}[\varphi_X^{-1}(A)]$$

,  $\forall A \in \mathcal{B}^T$

se denomina la ley o distribución del proceso  $X$ .

**Definición 3** Existe un conjunto  $F \subset T$ ,  $F$  es finito y se define a  $\mathbb{P}_F = \mathbb{P} \circ \{\varphi_X^F\}^{-1}$ , entonces se define a  $\{\mathbb{P}_F\}_{F \subset T, F \text{ es finito}}$  como la familia de la distribución finito dimensional del proceso  $X$

Entonces si consideramos a un conjunto  $F \subset L$ , tenemos

$$\begin{aligned}\mathbb{P}_L \circ \{\pi_F^L\}^{-1} \\ \mathbb{P} \circ \{\varphi_X^L\}^{-1} \circ \{\pi_F^L\}^{-1} \\ \mathbb{P} \circ (\pi_F^L \circ \varphi_X^L)^{-1} \\ \mathbb{P} \circ (\varphi_X^F)^{-1} \\ \mathbb{P}_F\end{aligned}$$

Es decir se cumple el principio de la consistencia

$$\mathbb{P}_L \circ \{\pi_F^L\}^{-1} = \mathbb{P}_F$$

Por lo tanto  $\mathbb{P}_F$  es una medida de probabilidad que cumple con la condicion de consistencia en  $\mathcal{B}^F$

**Definición 4** Dos procesos  $Y$  e  $X$ , se dice que son equivalentes si tiene la misma distribución entonces decimos que  $Y$  es una versión del proceso  $X$  si  $\mathbb{P}(X_t = Y_t) = 1 \forall t \in T$ . Entonces  $Y$  e  $X$  decimos que son procesos indistinguibles si  $\mathbb{P}(X_t = Y_t, \forall t \in T) = 1$

Para estructurar la definición anterior decimos si para cada  $k \in T$  entonces  $X_k = Y_k$  son equivalentes con probabilidad 1, entonces decimos que  $Y$  es un versión de  $X$  si  $\mathbb{P}(X_k = Y_k) = 1 \forall k \in T$  es decir  $X_t = Y_t$  c.s.  $\forall t \in T$  entonces si para cada  $t_1, t_2, t_3, \dots, t_n$ , en donde  $t_1 < t_2 < t_3 < \dots < t_n$ ,  $\forall n \in \mathbb{N}$ , satisface:

$$(X_{t_1}, X_{t_2}, X_{t_3}, \dots, X_{t_n}) = (Y_{t_1}, Y_{t_2}, Y_{t_3}, \dots, Y_{t_n})$$

c.s.

Implica que ambos procesos tienen las mismas distribuciones es decir la familia de la distribución finito dimensional del proceso  $\{Y_t\}_{t \in T}$  coincide con la familia de las distribuciones finito dimensional  $\{X_t\}_{t \in T}$ , entonces  $X$  e  $Y$  tiene la misma distribución como conclusión son equivalentes.

### 2.3.2. Propiedades de los procesos Gaussianos

A partir del teorema de la construcción de kolmogorov (teorema 2.3.1), se desea obtener el proceso estocástico  $X$  tal que su familia de distribuciones finito dimensionales del proceso coincidan con la familia de distribuciones finito dimensionales dadas apriori, es decir un proceso estocástico, con un espacio de probabilidad  $(\Omega, \mathfrak{F}, \mathbb{P})$ , donde  $\Omega$  es el espacio muestral no vacío,  $\mathfrak{F}$  es un sigma de algebra sub conjunto del espacio muestral no vacío y  $\mathbb{P}$  la medida de probabilidad una medida en  $\mathfrak{F}$ , conjunto de tiempo  $T$  y espacio de estados  $\mathbb{R}$ .

**Teorema 2.3.1** (Construcción de Kolmogorov). Supóngase que para cada  $F \subset T$ , con  $F$  finito, existe una medida de probabilidad  $\mathbb{P}_S$  sobre  $\mathcal{B}^S$  tal que cumple la condición de consistencia, entonces existe una medida de probabilidad  $\mathbb{P}$  sobre  $\mathcal{B}^T$  tal que:

$$\mathbb{P}_S = \mathbb{P} \circ [\varrho_S]^{-1}, \forall S \subset T, S \text{ finito.}$$

En particular, el proceso canónico  $(\mathbb{R}_T, \mathcal{B}^T, \mathbb{P}, [\varrho_t]_{t \in T})$  satisface

$$\mathbb{P} \circ [\varphi_\varrho^S]^{-1} = \mathbb{P} \circ [\varrho^S]^{-1} = \mathbb{P}_S, \forall S \subset T, S \text{ finito.}$$

Por lo tanto la familia de distribuciones finito dimensionales  $[\mathbb{P}_S]_{S \subset T, F \text{ finito}}$  dadas apriori coincide con la familia de distribución finito dimensionales del proceso canónico  $[\varrho_t]_{t \in T}$ .

Como consecuencia de dicho teorema las distribuciones finito dimensionales de un proceso estocástico determina la distribución del proceso. Siendo este teorema de gran importancia para construir procesos estocásticos, específicamente los procesos gaussianos

**Proposición 2.3.2** *Un Proceso estocástico  $[X_t]_{t \in T}$  es llamado proceso gaussiano si solo si todas las distribuciones finito dimensionales son distribuciones normales, es decir si solo si, el vector aleatorio  $(X_{t_1}, X_{t_2}, X_{t_3}, \dots, X_{t_n})$  es gaussiano para toda  $(t_1, t_2, t_3, \dots, t_n) \in T$ , con  $n \in \mathbb{N}$*

La siguiente proposición sigue de la proposición 2.3.2.

**Proposición 2.3.3** *El proceso  $[X_t]_{t \in T}$  se dice que es gaussiano si solo si todas las combinaciones lineales del proceso tienen distribución normal, es decir si solo si la variable aleatoria:  $a_1 X_{t_1} + a_2 X_{t_2} + a_3 X_{t_3} + \dots + a_n X_{t_n}$  es gaussiana para toda  $(t_1, t_2, t_3, \dots, t_n) \in T$  y  $a_1, a_2, a_3, \dots, a_n \in \mathbb{N}$ .*

La distribución del proceso  $[X_t]_{t \in T}$  está determinado por la función de esperanza  $m$  y el operador estadístico de covarianza  $K$  basado en el teorema de construcción de kolmogorov 2.3.1.

Por lo tanto según el teorema las distribuciones finito dimensionales del proceso  $[X_t]_{t \in T}$  definen la distribución del proceso  $[X_t]_{t \in T}$ , por lo cual la función de esperanza  $m_X(t) = E(X_t)$  y operador de covarianzas  $K_X(s, t) = cov(X_t, X_s)$  con  $s$  y  $t \in T$  determinan la distribución del proceso  $[X_t]_{t \in T}$

**Teorema 2.3.4** *Sea  $[X_t]_{t \in T}$  un proceso gaussiano, entonces su matriz de covarianzas  $K$  es:*

1. *Simétrica*
2. *Postiva definida, es decir, para cada,  $(t_1, t_2, t_3, \dots, t_n) \in T$  y  $a_1, a_2, a_3, \dots, a_n \in \mathbb{N}$ . Se cumple que:  $\sum_{i,j=1}^n a_i a_j K(t_i t_j) \geq 0$ . Recíprocamente, para cada función  $m : T \rightarrow \mathbb{R}$  y para cada función  $K : T^2 \rightarrow (\mathbb{R}, \text{simétrica y positiva definida})$ , existe un proceso gaussiano  $X$  tal que  $m_X \equiv m$  y  $K_X \equiv K$ . (Demostración en Tudor, 2002)*

**Teorema 2.3.5** *Sea  $[X_t]_{t \in T}$  un proceso gaussiano centrado con función de covarianzas  $K$ . Supongase que existen constantes  $\alpha, c > 0$  tales que:  $K(t, t) + K(s, s) - 2K(s, t) \leq c|t - s|^\alpha, \forall s, t \in T$ . Entonces existe una versión continua  $\hat{X}$  del proceso  $X$  tal que para todo  $0 < \gamma < \frac{\alpha}{2}$ , las trayectorias de  $\hat{X}$  son continuas  $\gamma$ -Holderianas c.s., (Azucena Campillo, 2013)*

*Demostración..* Sea  $\exists c, \alpha > 0$  Entonces,  $var(\omega_t - \omega_s) \leq c|t - s|^\alpha$

$$var(\omega_t - \omega_s) \leq c|t - s|^\alpha$$

$$\mathbb{E}(\omega_t - \omega_s)^2 \leq c|t - s|^\alpha$$

$$\mathbb{E}(\omega_t^2 - 2\omega_t\omega_s + \omega_s^2) \leq c|t - s|^\alpha$$

$$\mathbb{E}(\omega_t^2) - 2\mathbb{E}(\omega_t\omega_s) + \mathbb{E}(\omega_s^2) \leq c|t - s|^\alpha$$

$$\begin{aligned} cov(\omega_t, \omega_s) &= \mathbb{E}(\omega_t \omega_s) - \mathbb{E}(\omega_t) \mathbb{E}(\omega_s) \\ var(\omega_t) - 2cov(\omega_t \omega_s) + var(\omega_s) &\leq c|t - s|^\alpha \end{aligned}$$

$$K_\omega(t, t) - 2K_\omega(t, s) + K_\omega(s, s) \leq c|t - s|^\alpha$$

Si  $Y \sim N(0, \sigma^2)$ , Si consideramos  $\mathbb{E}(Y^{2m}) = \frac{2m!}{2^m m!} \sigma^{2m} \forall m \in \mathbb{N}$

Entonces si aplicamos:

$$\mathbb{E}(|\omega_t - \omega_s|^{2m}) = \frac{2m!}{2^m m!} (var|\omega_t - \omega_s|)^m$$

$$\mathbb{E}(|\omega_t - \omega_s|^{2m}) = \frac{2m!}{2^m m!} (\mathbb{E}(\omega_t - \omega_s)^2)^m$$

$$\mathbb{E}(|\omega_t - \omega_s|^{2m}) \leq \mathbb{E}^m |\omega_t - \omega_s|^{2m}$$

$$\mathbb{E}(|\omega_t - \omega_s|^{2m}) \leq c^m |t - s|^{\alpha m} \quad (2.1)$$

Si  $\alpha m > 1$ , entonces  $\alpha m - 1 = b$  y  $2m = a$ , si reemplazamos en la ecuación anterior tenemos:

$$\mathbb{E}(|\omega_t - \omega_s|^a) \leq c^m |t - s|^{b+1} \quad (2.2)$$

Por lo cual para  $m$  cumple las condiciones de continuidad de kolmogorov, entonces existe una versión continua con trayectorias  $\gamma$  Holderianas para cada  $0 < \gamma < \frac{b}{a}$

$$\text{Donde: } \lim_{m \rightarrow \infty} \left( \frac{\alpha m - 1}{2m} \right) = \lim_{m \rightarrow \infty} \left( \frac{\alpha}{2} \right) = \frac{\alpha}{2}$$

Aplicando L' Hospital, tenemos:

Entonces existe una versión continua cuyas trayectorias son  $\gamma$  Holderianas c.s para toda  $0 < \gamma < \frac{\alpha}{2}$

Si consideramos la regularidad de las trayectorias:

**Definición 5** Si una función  $X : T \rightarrow \mathbb{R}$ , se dice que es  $\gamma$ -Holder con  $0 < \gamma < 1$  si existe una constante  $C_\gamma$  tal que:

$$|X_t - X_s| \leq C_\gamma |t - s|^\gamma$$

$\forall s, t \in T$ .

**Teorema 2.3.6** *Criterio de continuidad de kolmogorov. Sea  $\{X_t\}_{t \in T}$  un proceso estocástico que toma valores en  $\mathbb{R}$ . Supóngase que existen constantes  $a, b, c > 0$ , tales que  $\forall s, t \in T$ , se cumple:*

$$\mathbb{E}(|X_t - X_s|^a) \leq c|t - s|^{1+b}$$

*Entonces, para cada  $0 < \gamma < \frac{b}{a}$ , existe una versión continua  $\{\tilde{X}_t\}_{t \in T}$  del proceso  $\{X_t\}_{t \in T}$  y una variable aleatoria positiva  $Y_\gamma$  que satisface:*

1.  $\mathbb{E}(Y_\gamma^a) < \infty$
2.  $|\tilde{X}_t - \tilde{X}_s| \leq Y_\gamma |t - s|^\gamma \forall s, t \in T$

**Proposición 2.3.7** *Un proceso gaussiano  $\{X_t\}_{t \in T}$  es estacionario si solo si su función de esperanza  $m_X$  es constante y su operador de covarianza  $K$  cumple la igualdad  $K(s, t) = K(s + h, t + h)$  para cada  $t, s \in T$  y todo  $h > 0$  (Azucena Campillo, 2013).*

Si el proceso  $\{X_t\}_{t \in T}$  es un proceso estacionario si  $h > 0$ , entonces  $\{X_{t+h}\}_{t \in T}$  converge en distribución a  $\{X_t\}_{t \in T}$  para todo  $h > 0$ . Entonces se cumple que:

$\mathbb{E}(X_{t+h}) = \mathbb{E}(X_t)$ ,  $\forall t \in T$  y  $h > 0$  por lo tanto la función esperanza  $\mathbb{E}(X)$  es constante.

Entonces las distribuciones finito dimensionales del proceso  $\{X_t\}_{t \in T}$  coincide con las distribuciones finito dimensionales del proceso  $\{X_{t+h}\}_{t \in T} \forall t \in T$  y  $h > 0$  por lo cual se cumple:

$(X_{t+h}, X_{s+h})$  converge en distribución a  $(X_t, X_s)$ , entonces:

$$Cov(X_{t+h}, X_{s+h}) = cov(X_t, X_s)$$

$$K_X(t + h, s + h) = K_X(t, s)$$

,

### 2.3.3. Propiedades generales de distribuciones de procesos

1. Un proceso  $X$  se dice que tiene incrementos independientes si  $\forall t_i \in T$ ,  $i=1,2,3,\dots,n$ , con  $0 < t_1 < t_2 < t_3 < \dots < t_{n-1} < t_n$ , la variables aleatorias  $X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$ , son independientes.

2. Un proceso  $X$  se dice estacionario si para toda  $\forall t_i \in T$ ,  $i=1,2,3,\dots,n$ , se tiene que:  
 $(X_{t_1+h}, X_{t_2+h}, X_{t_3+h}, \dots, X_{t_n+h})$  converge en distribución a  $(X_{t_1}, X_{t_2}, X_{t_3}, \dots, X_{t_n})$   
 con  $h > 0$  también denotado por  $\{X_{t+h}\}_{t \in T}$  converge en distribución a  $\{X_t\}_{t \in T}$
  
3. Un proceso  $X$  se dice que tiene incrementos estacionarios si las distribuciones de los incrementos  $X_{t+h} - X_h$ , con  $t \in T$ , no depende de  $h > 0$ .



### 2.3.4. Movimiento Browniano

El movimiento Browniano es un fenómeno que debe su nombre al botánico británico Robert Brown quien en 1827 observó el comportamiento constante y aleatorio de las partículas de polen en un fluido y encontró también que cualquier objeto sea orgánico o inorgánico mostraba este mismo comportamiento entonces concluyó que estaba relacionado con el agua pero no llegó a explicar las causas de su movimiento, luego en el año de 1905 el segundo artículo publicado por Albert Einstein sobre el movimiento de las partículas pequeñas en un líquido estacionario en el que predecía un fenómeno que había sido detectado hace más de setenta años, pero su aporte a la ciencia fue la construcción de un modelo estadístico basada matemáticamente en trabajos anteriores de termodinámica y mecánica estadística con el que predecía el movimiento de moléculas suspendidas en un líquido, dicho modelo descriptivo de Einstein para el movimiento browniano es popularmente conocido hoy en día como movimiento aleatorio. Pero En 1900, L. Bachelier introdujo el proceso para modelar las fluctuaciones financieras. Por otro lado fue Norberth Wiener el primero en dar una rigurosidad en la formulación matemática del movimiento en 1923

Sea  $X = \{X_t\}_{t \in T}$  es un proceso gaussiano, entonces el operador estadístico de covarianza es:

$$K_\omega(s, t) = \min\{s, t\} \quad (2.3)$$

El matemático Norberth Wiener desarrollo matemáticamente desde el punto de vista de la probabilidad el movimiento browniano considerando los incrementos del proceso, por lo tanto para fines de demostración se considera un movimiento Browniano centrado a un proceso con función de esperanza  $m_\omega$  entonces llamaremos un movimiento Browniano estándar. Es importante mencionar su naturaleza autosimilar, entendiéndolo a un proceso autosimilar si el movimiento Browniano  $\omega^H$  con  $H \in \langle 0, 1 \rangle$  implica la existencia de dos procesos  $\{\omega_{at}^H\}_{t \in T}$  y  $\{a^H \omega_t^H\}_{t \in T}$  tienen la misma distribución para toda  $a > 0$  (en la sección movimiento browniano fraccional se desarrollará con más amplitud este concepto).

En la cual se cumplen las siguientes propiedades:

1.  $\omega_t \sim N(0, t)$
2.  $\omega_0 = 0$  casi seguramente
3. Para  $s \leq t$  entonces  $\omega_t - \omega_s \stackrel{D}{=} \omega_{t-s} \sim N(0, t - s)$ .
4.  $\omega$  es un proceso con incrementos estacionarios e independientes.

5.  $\omega$  es un proceso no estacionario.
6.  $\omega$  es un proceso con trayectorias continuas  $\gamma$  holderianas para todo  $0 < \gamma < \frac{1}{2}$ .

Demostración

1. Para  $\{\omega_t\}_{t \in T}$  es un movimiento Browniano estandar con  $\omega_t \sim N(0, t)$  si

Si es un proceso autosimilar  $\omega_t = t^H \omega_1$ ,  $\omega_1 \sim N(0, 1)$ , lo daremos por definición (en la sección de movimiento browniano fraccional desarrollaremos la demostración)

$$\mathbb{E}[\omega_t] = \mathbb{E}[t^H \omega_1] = t^H \mathbb{E}[\omega_1]$$

$$t^H \mathbb{E}[\omega_1] = t^H(0)$$

$$\mathbb{E}[\omega_t] = 0$$

$$m_\omega(t) = 0$$

$$\text{var}(\omega_t) = \mathbb{E}[\omega_t]^2 = \mathbb{E}[t^H \omega_1]^2$$

$$t^{2H} \mathbb{E}[\omega_1]^2 = t^{2H}(1)$$

con  $H = \frac{1}{2}$ , entonces  $\text{var}(\omega_t) = t$

$$\text{var}(\omega_t) = K_\omega(t, t) = t$$

Entonces

$$\omega_t \sim N(0, t)$$

2. A partir de 1. Si  $\omega_t \sim N(0, t)$  entonces  $\omega_0 \sim N(0, 0)$

$$\mathbb{E}[\omega_0] = 0$$

$$\text{var}(\omega_0) = K_\omega(0, 0) = 0$$

Si  $\omega_0 \sim N(0, 0)$  la cual es denominada la distribución de Delta Dirac centrada en cero por lo cual si consideramos un espacio de medida  $(\Omega, \mathfrak{F}, P)$  donde

$\Omega \neq \emptyset$  y  $\mathfrak{S} \subset P(\Omega)$  una clase  $\mathfrak{S}$  subconjunto de  $\Omega$  no vacío,  $P : \mathfrak{S} \rightarrow [0, \infty)$   
Sabiendo que  $P$  es una medida en  $\mathfrak{S}$ .

$$P(\omega) = \begin{cases} 1 & \text{si } 0 \in \omega_0 \\ 0 & \text{si } 0 \notin \omega_0 \end{cases}$$

Entonces la medida está centrada en 0 es denotada por  $\delta_0$

$P(\omega_0 = 0) = 1$  convergencia casi ciertamente

3. Por la proposición

**Proposición 2.3.8** *El proceso gaussiano  $\{X_t\}_{t \in T}$  se dice que es gaussiano si solo si todas las combinaciones lineales del proceso tienen distribución normal, es decir si solo si la variable aleatoria  $a_1x_{t_1} + \dots + a_nx_{t_n}$  es gaussiana para toda  $t_1, \dots, t_n \in T$  y  $a_1 \dots a_n \in \mathbb{R}$ , con  $n \in \mathbb{N}$  (Tudor, 2002)*

Entonces  $\omega_t$  es gaussiano debido a la proposición 2.3.2 dicha proposición surge como resultado del teorema de construcción de kolmogorov y por la proposición 2.3.3 la distribución del proceso  $\omega_t - \omega_s$  es gaussiana, por lo tanto  $\omega_t - \omega_s$  tienen distribución normal, entonces faltaría solamente demostrar que  $\omega_{t-s} \sim N(0, t-s)$ .

Entonces por ser un proceso Autosímil se cumple:  $\omega_{t-s} = (t-s)^H \omega_1$  con  $H \in (0, 1)$  (se desarrollará con mas profundidad en la sección del movimiento Browniano Fraccional).

$$\mathbb{E}[\omega_{t-s}] = \mathbb{E}[(t-s)^H \omega_1] = (t-s)^H \mathbb{E}[\omega_1]$$

Si  $\omega_1 \sim N(0, 1)$ .

$$\mathbb{E}[\omega_{t-s}] = (t-s)^H (0) = 0$$

$$\mathbb{E}[\omega_{t-s}]^2 = \mathbb{E}[(t-s)^H \omega_1]^2 = (t-s)^{2H} \mathbb{E}[\omega_1]^2 = (t-s)^{2H} (1) = (t-s)^{2H}$$

con  $H = \frac{1}{2}$ .

$$\mathbb{E}[\omega_{t-s}]^2 = (t-s) = \text{var}(\omega_{t-s})$$

Por lo tanto  $\omega_{t-s} \sim N(0, t-s)$ .

Definido de un manera alternativa:

$$\mathbb{E}[\omega_t - \omega_s] = \mathbb{E}[\omega_t] - \mathbb{E}[\omega_s] = 0$$

$$\mathbb{E}[\omega_t - \omega_s] = 0$$

$$var(\omega_t - \omega_s) = var(\omega_t) - var(\omega_s)$$

$$var(\omega_t) - var(\omega_s) = t - s$$

$$\mathbb{E}[\omega_t - \omega_s]^2 = \mathbb{E}[\omega_t^2 - 2\omega_t\omega_s + \omega_s^2]$$

$$\mathbb{E}[\omega_t - \omega_s]^2 = \mathbb{E}[\omega_t^2] - 2\mathbb{E}[\omega_t\omega_s] + \mathbb{E}[\omega_s^2]$$

Sabiendo que  $K_\omega(t, s) = cov(\omega_t, \omega_s)$

$$cov(\omega_t, \omega_s) = \mathbb{E}[\omega_t\omega_s] - \mathbb{E}[\omega_t]\mathbb{E}[\omega_s]$$

Si  $\omega_t \sim N(0, t)$  entonces

$$cov(\omega_t, \omega_s) = \mathbb{E}[\omega_t\omega_s]$$

si aplicamos la propiedad 2, entonces:

$$\mathbb{E}[\omega_t - \omega_s]^2 = t - 2s + s = t - s$$

$$\omega_t - \omega_s \sim N(0, t - s)$$

Por lo tanto  $\omega_t - \omega_s$  converge en distribución finito dimensional a  $\omega_{t-s} \sim N(0, t-s)$

4. A partir de 3 tengo  $(\omega_{t+h} - \omega_h)$  converge en distribución a  $\omega_t \forall h > 0$  y  $t \in T$ .

Por lo tanto si:

$var(\omega_{t+h} - \omega_h)$  no tiene que estar en función de h

$$\begin{aligned} & \mathbb{E}[\omega_{t+h}]^2 - 2\mathbb{E}[\omega_{t+h}\omega_h] + \mathbb{E}[\omega_h]^2 \\ & var[\omega_{t+h}] - 2cov[\omega_{t+h}, \omega_h] + var[\omega_h] = t + h - 2h + h = t \end{aligned}$$

no depende de h, por lo tanto decimos que el proceso  $\{\omega_t\}_{t \in T}$  tiene incrementos estacionarios.

Ahora si consideramos para cualesquiera  $s_1, s_2, t_1, t_2 \in T$  y  $0 \leq s_1 < s_2 < t_1 < t_2$ , entonces el vector  $(\omega_{t_2} - \omega_{t_1}, \omega_{s_2} - \omega_{s_1})$  es gaussiano pues son combinaciones lineales del movimiento Browniano  $\{\omega_t\}_{t \in T}$ . segun la proposición siguiente:

**Proposición 2.3.9** *Un vector aleatorio  $X = (X_1, \dots, X_n)$  es un vector gaussiano si solo si todas las combinaciones lineales de sus coordenadas  $a_1X_1 + a_2X_2 + a_3X_3 + \dots + a_nX_n$  tiene distribución gaussiana en  $\mathbb{R}$ , para toda  $a = (a_1, a_2, a_3, \dots, a_n) \in \mathbb{R}^n$  (Tudor, 2002).*

Por otro lado para comprobar la independencia de los incrementos de  $(\omega_{t_2} - \omega_{t_1}, \omega_{s_2} - \omega_{s_1})$  según la proposición

**Proposición 2.3.10** *Si  $x$  e  $y$  son variables aleatorias independientes e integrables entonces (Barry James, 1981):*

$$\mathbb{E}(xy) = \mathbb{E}(x)\mathbb{E}(y)$$

Si  $X$  e  $Y$  son independientes  $\Rightarrow f(x, y) = f_1(x) \cdot f_2(y)$

$$\mathbb{E}(xy) = \iint xy(x, y) dx dy$$

$$\mathbb{E}(xy) = \int_{R_y} \int_{R_x} xy f_1(x) f_2(y) dx dy$$

$$\mathbb{E}(xy) = \int_{R_y} y \left[ \int_{R_x} x f_1(x) dx \right] \cdot f_2(y) dy$$

Si  $\mathbb{E}(x) = \int_{R_x} x f_1(x) dx$ , entonces

$$\mathbb{E}(xy) = \int_{R_y} y \mathbb{E}(x) \cdot f_2(y) dy$$

$$\mathbb{E}(xy) = \mathbb{E}(x) \int_{R_y} y f_2(y) dy$$

$$\mathbb{E}(xy) = \mathbb{E}(x)\mathbb{E}(y)$$

**Proposición 2.3.11** Sea  $(X, Y)$  un vector gaussiano, entonces  $X$  e  $Y$  son independientes si y solo si  $cov(X, Y) = 0$  (Barry James, 1981).

Entonces si aplico la proposición 2.3.10 y 2.3.11 se demuestra:

$$cov(\omega_{t_2} - \omega_{t_1}, \omega_{s_2} - \omega_{s_1}) = cov(\omega_{t_2}, \omega_{s_2}) - cov(\omega_{t_2}, \omega_{s_1}) - cov(\omega_{t_1}, \omega_{s_2}) + cov(\omega_{t_1}, \omega_{s_1})$$

Si utilizamos la ecuación 2.3, se tiene:

$$K_\omega(t_2 - t_1, s_2 - s_1) = K_\omega(t_2, s_2) - K_\omega(t_2, s_1) - K_\omega(t_1, s_2) + K_\omega(t_1, s_1)$$

$$K_\omega(t_2 - t_1, s_2 - s_1) = 0$$

$$cov(\omega_{t_2} - \omega_{t_1}, \omega_{s_2} - \omega_{s_1}) = 0$$

Entonces el proceso  $\{\omega_t\}_{t \in T}$  tiene incrementos estacionarios e independientes.

5. Para demostrar que el proceso  $\{\omega_t\}_{t \in T}$  no es estacionario basta con probar que  $K_\omega(s + h, t + h) \neq K_\omega(s, t) \forall h > 0$ , según la proposición 2.3.7.

Entonces:

$$cov(\omega_{t+h}, \omega_{s+h}) \neq cov(\omega_t, \omega_s)$$

$$cov(\omega_{t+h}, \omega_{s+h}) = \mathbb{E}(\omega_{t+h}, \omega_{s+h}) - \mathbb{E}(\omega_{t+h})\mathbb{E}(\omega_{s+h})$$

$$cov(\omega_{t+h}, \omega_{s+h}) = \mathbb{E}(\omega_{t+h}, \omega_{s+h}) - \mathbb{E}(\omega_t + \omega_h)\mathbb{E}(\omega_s + \omega_h)$$

$$cov(\omega_{t+h}, \omega_{s+h}) = \mathbb{E}(\omega_{t+h}, \omega_{s+h}) - [\mathbb{E}(\omega_t) + \mathbb{E}(\omega_h)][\mathbb{E}(\omega_s) + \mathbb{E}(\omega_h)]$$

$$cov(\omega_{t+h}, \omega_{s+h}) = \mathbb{E}(\omega_{t+h}, \omega_{s+h})$$

Por otro lado:

$$cov(\omega_t, \omega_s) = \mathbb{E}(\omega_t \omega_s)$$

$$K_\omega(s+h, t+h) \neq K_\omega(s, t)$$

$$\forall h > 0$$

Entonces ambas covarianzas son distintas por lo cual el proceso no es estacionario

$$\text{cov}(\omega_{t+h}, \omega_{s+h}) \neq \text{cov}(\omega_t, \omega_s)$$

$\{\omega_t\}_{t \in T}$  no es estacionaria,  $\{\omega_{t+h}\}_{t \in T}$  no converge en distribución finito dimensional a  $\{\omega_t\}_{t \in T} \forall h > 0$ .

6. Para demostrar la existencia de trayectorias continuas se hará uso del siguiente teorema derivado del teorema 2.3.5:

**Teorema 2.3.12** Sea  $\{\omega_t\}_{t \in T}$  un proceso gaussiano centrado con función de covarianza  $K$  Suponemos que existen constantes  $\alpha, c > 0$  tales que:

$$\text{cov}(\omega_t, \omega_t) + \text{cov}(\omega_s, \omega_s) - 2\text{cov}(\omega_t, \omega_s) \leq c|t-s|^\alpha \quad \forall s, t \in T$$

Sabiendo que:  $K_\omega(t, t) = \text{cov}(\omega_t, \omega_t)$ ,  $K_\omega(s, s) = \text{cov}(\omega_s, \omega_s)$ ,  $K_\omega(t, s) = \text{cov}(\omega_t, \omega_s)$ , Se puede expresar:

$$K_\omega(t, t) + K_\omega(s, s) - 2K_\omega(t, s) \leq c|t-s|^\alpha$$

$$, \forall s, t \in T.$$

Entonces existe una versión continua  $\tilde{\omega}$  del proceso  $\omega$  para toda  $0 < \gamma < \frac{\alpha}{2}$  las trayectorias son  $\gamma$  holderianas c.s.

Entonces sea  $t, s \in T$  si debido a la propiedad 1  $\omega_t \sim N(0, t)$  y  $\omega_s \sim N(0, s)$

$\omega_t - \omega_s$  converge en distribución a  $\omega_{t-s} \sim N(0, t-s)$

Si  $\text{var}(\omega_t - \omega_s) \leq c|t-s|^\alpha \quad \forall s, t \in T$

$$\mathbb{E}[\omega_t - \omega_s]^2 \leq c|t-s|^\alpha$$

$$\mathbb{E}[\omega_t^2 + \omega_s^2 - 2(\omega_t \omega_s)] \leq c|t-s|^\alpha$$

$$\mathbb{E}[\omega_t^2] + \mathbb{E}[\omega_s^2] - 2\mathbb{E}(\omega_t \omega_s) \leq c|t-s|^\alpha$$

$$\text{var}(\omega_t) - 2\text{cov}(\omega_t \omega_s) + \text{var}(\omega_s) \leq c|t-s|^\alpha$$

$$t - 2k_\omega(t, s) + s \leq c|t-s|^\alpha$$

Aplicando la ecuación 2.3, tenemos:

$$t - 2s + s \leq c|t - s|^\alpha$$

$t - s \leq c|t - s|^\alpha$  , si  $c = 1$  y  $\alpha = 1$ , entonces por el teorema 2.3.12 se garantiza que existen trayectorias continuas  $\tilde{\omega}$  del proceso  $\omega$  las cuales son  $\gamma$  holderianas para cada  $0 < \gamma < \frac{1}{2}$



### 2.3.5. Movimiento Browniano Fraccional

El movimiento Browniano Fraccional fue analizado por primera vez por el matemático ruso Nikolayevich kolmogorov en 1940 y el matemático polonés Benoit Mandelbrot en 1965 para ser desarrollado e investigado con mayor interés por Mandelbrot y Van Ness en 1968 considerando la dependencia de largo plazo de los incrementos de un proceso, dicho modelo es denominado multifractal también como procesos estocásticos multifractales. Al igual que el movimiento browniano es un proceso autosimilar pero en este caso con índice de similaridad  $H \neq \frac{1}{2}$  a diferencia del browniano donde  $H = \frac{1}{2}$  como se desarrollara mas adelante.

Entonces el movimiento Browniano estándar denotado por  $B^H$  es un proceso gaussiano centrado con covarianza

$$K_H(t, s) = \frac{1}{2}(|s|^{2H} + |t|^{2H} - |t - s|^{2H}) \quad (2.4)$$

$\forall t, s \in T = \mathbb{R}^+$

En donde  $H \in \langle 0, 1 \rangle$  y  $H$  es llamado el coeficiente de hurts.

Análisis

Si consideramos la función de covarianza, entonces

$$K_H(t, s) = \frac{1}{2}(|s|^{2H} + |t|^{2H} - |t - s|^{2H})$$

, para  $H = \frac{1}{2}$

$$K_{\frac{1}{2}}(t, s) = \frac{1}{2}(|s| + |t| - |t - s|)$$

$$K_{\frac{1}{2}}(t, s) = \min\{t, s\}$$

$\forall t, s \in T$ , nos da la ecuación 2.3 es decir obtenemos un movimiento Browniano estandar  $\omega_t$

Como segundo paso para probar la no estacionariedad del movimiento Browniano Fraccional sean

$$K_H(s + h, t + h) = \frac{1}{2}(|(t + h)|^{2H} + |(s + h)|^{2H} - |t - s|^{2H})$$

$$K_H(s + h, t + h) \neq \frac{1}{2}(|(t)|^{2H} + |(s)|^{2H} - |t - s|^{2H})$$

$$K_H(s + h, t + h) \neq K_H(s, t)$$

Debido a no cumplir con la siguiente proposición que sigue a la proposición 2.3.7

**Proposición 2.3.13** *Un proceso gaussiano  $\{X_t\}_{t \in T}$  es estacionario si solo si su función de esperanza  $m_X$  es constante y su operador de covarianza  $K$  cumple la igualdad  $K_H(s+h, t+h) = K_H(s, t)$  para cada  $t, s \in T$  y para  $h > 0$ .*

El proceso sería estacionario si  $\{\omega_{t+h}\}_{t \in T}$  converge en distribución a  $\{\omega_t\}_{t \in T}$   $\forall h > 0$

En este caso el movimiento Browniano Fraccionario cumple que  $\{\omega_{t+h}\}_{t \in T}$  no converge en distribución a  $\{\omega_t\}_{t \in T} \forall h > 0$ . El movimiento Browniano Fraccionario es no estacionario.

**Proposición 2.3.14**

*El proceso  $B^H$  tiene incrementos estacionarios*

*Dichos incrementos son independientes si  $H = \frac{1}{2}$  es decir es un movimiento Browniano estandar.*

Demostracion

1. Sea

$$B_{t+h}^H \sim N(0, |t+h|^{2H})$$

$$B_h^H \sim N(0, |h|^{2H})$$

$\forall h > 0$

Entonces el operador estadístico de la varianza es

$$var(B_{t+h}^H - B_h^H) = \mathbb{E}[B_{t+h}^H - B_h^H]^2$$

$$\mathbb{E}[B_{t+h}^H - B_h^H]^2 = \mathbb{E}[(B_{t+h}^H)^{2H} - 2(B_{t+h}^H)(B_h^H) + (B_h^H)^{2H}]$$

$$var(B_{t+h}^H - B_h^H) = |t+h|^{2H} - 2cov(B_{t+h}^H, B_h^H) + |h|^{2H}$$

$$var(B_{t+h}^H - B_h^H) = |t+h|^{2H} - 2K_B^H(t+h, h) + |h|^{2H}$$

$$var(B_{t+h}^H - B_h^H) = |t+h|^{2H} - 2[\frac{1}{2}(|h|^{2H} + |t+h|^{2H} - |h-t-h|^{2H})] + |h|^{2H}$$

$$var(B_{t+h}^H - B_h^H) = |t+h|^{2H} - [|h|^{2H} + |h|^{2H} - |t|^{2H}]$$

$$var(B_{t+h}^H - B_h^H) = |t+h|^{2H} + |h|^{2H} - |h|^{2H} - |t+h|^{2H} + |t|^{2H}$$

$$\text{var}(B_{t+h}^H - B_h^H) = t^{2H}$$

Como la varianza no depende de  $h$  entonces el movimiento Browniano Fraccionario tiene incrementos estacionarios.

2. Si tenemos un vector formado por  $(B_{s_2}^H - B_{s_1}^H, B_{t_2}^H - B_{t_1}^H) \forall s_1, s_2, t_1, t_2 \in T$   $s_1 < s_2 < t_1 < t_2$  entonces es un vector gaussiano debido a que el resultado de la combinación lineal de variables gaussianas, para calcular la covarianza entre variables tengo  $\text{cov}(B_{s_2}^H - B_{s_1}^H, B_{t_2}^H - B_{t_1}^H)$ , entonces si desarrollamos:

$$\text{cov}(B_{s_2}^H - B_{s_1}^H, B_{t_2}^H - B_{t_1}^H) = \text{cov}(B_{s_2}^H B_{t_2}^H) - \text{cov}(B_{s_2}^H B_{t_1}^H) - \text{cov}(B_{s_1}^H B_{t_2}^H) + \text{cov}(B_{s_1}^H B_{t_1}^H)$$

$$K_B^H(s_2, t_2) - K_B^H(s_2, t_1) - K_B^H(s_1, t_2) + K_B^H(s_1, t_1)$$

$$\frac{1}{2}[|t_2|^{2H} + |s_2|^{2H} - |t_2 - s_2|^{2H}] - \frac{1}{2}[|t_1|^{2H} + |s_2|^{2H} - |t_1 - s_2|^{2H}] - \frac{1}{2}[|t_2|^{2H} + |s_1|^{2H} - |t_2 - s_1|^{2H}] + \frac{1}{2}[|t_1|^{2H} + |s_1|^{2H} - |t_1 - s_1|^{2H}]$$

$$\frac{1}{2}[|t_2 - s_2|^{2H}] + \frac{1}{2}[|t_1 - s_2|^{2H}] + \frac{1}{2}[|t_2 - s_1|^{2H}] - \frac{1}{2}[|t_1 - s_1|^{2H}]$$

$$\text{cov}(B_{s_2}^H - B_{s_1}^H, B_{t_2}^H - B_{t_1}^H) = \frac{1}{2}[|t_1 - s_2|^{2H} + |t_2 - s_1|^{2H} - |t_2 - s_2|^{2H} - |t_1 - s_1|^{2H}]$$

Según la proposición 2.3.11 como la  $\text{cov}(B_{s_2}^H - B_{s_1}^H, B_{t_2}^H - B_{t_1}^H)$  es diferente de cero entonces decimos que los incrementos no son independientes, por lo cual se concluye que los incrementos son estacionarios y dependientes. Si  $H = \frac{1}{2}$  entonces la covarianza es cero por lo cual se diría que es un movimiento Browniano estándar implicarían incrementos estacionarios e independientes.

Una de las propiedades más importantes del movimiento Browniano fraccional es la autosimilitud la cual se define como la invarianza de la distribución en tiempo y espacio debido a este fenómeno se le denomina como proceso autosimilar cuya manera de identificar en una serie es mediante el coeficiente de hurts o de autosimilitud.

Entonces para probar que el movimiento Browniano fraccional es un proceso autosimilar se hará uso de la siguiente proposición:

**Proposición 2.3.15** *Sea un proceso  $\{\omega_t\}_{t \in T}$  entonces decimos que dicho proceso es autosimilar con índice  $H > 0$  si para cada  $a > 0$ , existen procesos  $\{\omega_{at}\}_{t \in T}$  y  $\{a^H \omega_t\}_{t \in T}$  las cuales tienen la misma distribución también denotado por  $\{a^{-H} \omega_{at}\}_{t \in T}$  como también  $\{\omega_t\}_{t \in T}$  tienen la misma distribución.*

Entonces por la proposición 2.3.15, el movimiento Browniano fraccional  $B^H$  con  $H \in \langle 0, 1 \rangle$  implica la existencia de dos procesos  $\{B_{at}^H\}_{t \in T}$  y  $\{a^H B_t^H\}_{t \in T}$  las cuales segun la definición anterior tienen la misma distribución para toda  $a > 0$ . Es necesario para la simplicidad de la demostración que el proceso  $B^H$  sea un proceso gaussiano centrado por lo que tambien los procesos  $\{B_{at}^H\}_{t \in T}$  y  $\{a^H B_t^H\}_{t \in T}$  son gaussianos, entonces seria solamente necesario probar la igualdad de las funciones de covarianzas de ambos procesos para demostrar que tienen una misma distribución. Entonces:

Sea  $s, t \in T$  y  $a > 0$  entonces se tiene:

$$\begin{aligned}
cov(B_{at}^H, B_{st}^H) &= K_H(at, as) \\
&= \frac{1}{2} [|at|^{2H} + |as|^{2H} - |at - as|^{2H}] \\
&= \frac{1}{2} [a^{2H} |t|^{2H} + a^{2H} |s|^{2H} - a^{2H} |t - s|^{2H}] \\
&= \frac{1}{2} a^{2H} [|t|^{2H} + |s|^{2H} - |t - s|^{2H}] \\
&= a^{2H} \frac{1}{2} [|t|^{2H} + |s|^{2H} - |t - s|^{2H}] \\
&= a^{2H} cov(B_t^H, B_s^H) \\
cov(a^{2H} B_t^H, a^{2H} B_s^H) &= a^{2H} K_H(t, s) \\
a^{2H} K_H(t, s) &= a^{2H} \frac{1}{2} [|t|^{2H} + |s|^{2H} - |t - s|^{2H}] \\
cov(B_{at}^H, B_{as}^H) &= cov(a^H B_t^H, a^H B_s^H)
\end{aligned}$$

Por lo tanto se demuestra que es un proceso autosimilar.

Una propiedad importante del movimiento Browniano fraccionario es la regularidad de las trayectorias, entonces a partir de la proposición :

## Regularidad de las trayectorias

**Proposición 2.3.16** *El movimiento Browniano fraccionario con índice de autosimilaridad  $H \in \langle 0, 1 \rangle$ ,  $B^H$ , admite una versión continua  $\tilde{B}^H$  tal que para toda  $\gamma < H$ , las trayectorias del proceso  $\tilde{B}^H$  son  $\gamma$  Holderianas c.s.*

Entonces si desarrollamos :

$$var(B_t^H - B_s^H) \leq c |t - s|^\alpha \quad \forall \alpha, c > 0$$

$$\mathbb{E}[[B_t^H]^2 - 2B_t^H B_s^H + [B_s^H]^2] \leq c|t - s|^\alpha$$

$$\mathbb{E}[B_t^H]^2 - 2\mathbb{E}[B_t^H B_s^H] + \mathbb{E}[B_s^H]^2 \leq c|t - s|^\alpha$$

$$\text{var}(B_t^H) - 2\text{cov}[B_t^H, B_s^H] + \text{var}(B_s^H) \leq c|t - s|^\alpha$$

$$K_H(t, t) - 2K_H(t, s) + K_H(s, s) \leq c|t - s|^\alpha$$

Si sabemos que:

$$B_t^H \sim N(0, |t|^{2H})$$

$$B_s^H \sim N(0, |s|^{2H})$$

$$K_H(t, s) = \frac{1}{2}[|t|^{2H} + |s|^{2H} - |t - s|^{2H}]$$

Entonces:

$$K_H(t, t) - 2K_H(t, s) + K_H(s, s) \leq c|t - s|^\alpha$$

$$|t|^{2H} - 2\left[\frac{1}{2}[|t|^{2H} + |s|^{2H} - |t - s|^{2H}]\right] + |s|^{2H} \leq c|t - s|^\alpha$$

$$|t|^{2H} - |t|^{2H} - |s|^{2H} + |t - s|^{2H} + |s|^{2H} \leq c|t - s|^\alpha$$

$$|t - s|^{2H} \leq c|t - s|^\alpha$$

Entonces si  $c = 1$  y  $\alpha = 2H$  por el teorema 2.3.5 tengo que para cada  $0 < H < 1$  se da la existencia de una versión continua  $\tilde{B}_t^H$  para toda  $0 < \gamma < \frac{\alpha}{2}$ , entonces si  $\alpha = 2H$  tengo  $\gamma < H$  las trayectorias del proceso son  $\gamma$  Holderianas c.s.

Entonces el movimiento Browniano fraccional cumple las condiciones de continuidad de Kolmogorov decimos que tiene trayectorias continuas Holder c.s. con orden menor a  $H$  (Sottinen 2001) y debido a la propiedad de autosimilaridad es decir ser un proceso autosimilar entonces sus trayectorias son diferenciables en ninguna parte c.s.(Mandelbrot y Ness (1968)).

**Proposición 2.3.17** *Sea un proceso gaussiano  $\{X_t\}_{t \in T}$  autosimilar con índice de similaridad  $H$  con incrementos estacionarios, entonces:*

1.  $X_0 = 0$  c.s

2.  $0 < H < 1$  , entonces  $\exists s, t \in T \forall s \leq t$  con operador de covarianzas  $K$

$$K_H(t, s) = \frac{1}{2}[|t|^{2H} + |s|^{2H} - |t - s|^{2H}]var(X_1)$$

3.  $H = 1$ , entonces  $X_t = tX_1$

Entonces

1. Si aplicamos el principio de similaridad entonces,  $\{X_{at}\}_{t \in T}$  y  $\{a^H X_t\}_{t \in T}$  tienen que tener la misma distribución , entonces:

$$K_X(at, at) = cov(a^H X_t, a^H X_t)$$

$$\forall a > 0$$

$$var(X_{at}) = var(a^H X_t)$$

$$\mathbb{E}[X_{at}^2] = \mathbb{E}[a^H X_t^2]$$

$$\text{Si } t = 0$$

$$\mathbb{E}[X_0^2] = \mathbb{E}[a^H X_0^2]$$

$$\mathbb{E}[X_{at}^2] = a^{2H} \mathbb{E}[X_0^2]$$

Por autosimilitud

$$X_t = tX_1 \text{ entonces } X_{t0} = 0X_1 \text{ por lo tanto } X_0 = 0$$

$$\mathbb{E}[X_0^2] = a^{2H} \mathbb{E}[X_0^2]$$

$$\mathbb{E}[X_0^2] = a^{2H} \mathbb{E}[0]$$

Entonces:

$$\mathbb{E}[X_0^2] = 0 \text{ entonces } X_0 \sim N(0, 0) \text{ por lo tanto } X_0 \sim N(0, 0^{2H})$$

$$X_0 \sim N(0, 0) \text{ se concluye } X_0 = 0 \text{ c.s.}$$

2.  $0 < H < 1$

$$K_H(t, s) = \frac{1}{2} [|t|^{2H} + |s|^{2H} - |t - s|^{2H}] \text{var}(X_1)$$

Si aplicamos las condiciones de incrementos estacionarios y principio de autosimilaridad

Entonces

$$\mathbb{E}(X_1) = \mathbb{E}(X_2 - X_1)$$

Si  $X_2 = 2^H X_1$

$$\mathbb{E}(X_1) = \mathbb{E}[X_2] - \mathbb{E}[X_1]$$

$$\mathbb{E}(X_1) = \mathbb{E}[2^H X_1] - \mathbb{E}[X_1]$$

$$\mathbb{E}(X_1) = 2^H \mathbb{E}[X_1] - \mathbb{E}[X_1]$$

$$\mathbb{E}(X_1) = [2^H - 1] \mathbb{E}[X_1]$$

con  $H = 0$

$$\mathbb{E}(X_1) = [2^0 - 1] \mathbb{E}[X_1]$$

$$\mathbb{E}[X_1] = 0$$

Si por el principio de autosimilitud  $\{X_{at}^H\} \doteq \{a^H X_t^H\}$

$$\mathbb{E}[X_{at}] = \mathbb{E}[a^H X_t]$$

$$\mathbb{E}[X_{at}] = a^H \mathbb{E}[X_t]$$

Para  $t=1$

$$\mathbb{E}[X_a] = a^H \mathbb{E}[X_1]$$

$$\mathbb{E}[X_a] = 0$$

Entonces el proceso  $\{X_t\}_{t \in T}$  gaussiano es centrado en cero.

Por otro lado

$$\text{var}[X_t - X_s] = \mathbb{E}[X_t - X_s]^2$$

$$\mathbb{E}[X_t^2 - 2[X_t X_s] + X_s^2]$$

$$\mathbb{E}[X_t^2] - 2\mathbb{E}[X_t X_s] + \mathbb{E}[X_s^2]$$

$$\text{var}[X_t - X_s] = \mathbb{E}[X_t^2] - 2\mathbb{E}[X_t X_s] + \mathbb{E}[X_s^2]$$

$$2\mathbb{E}[X_t X_s] = \mathbb{E}[X_t^2] + \mathbb{E}[X_s^2] - \text{var}[X_t - X_s]$$

$$\mathbb{E}[X_t X_s] = \frac{1}{2}[\mathbb{E}[X_t^2] + \mathbb{E}[X_s^2] - \text{var}[X_t - X_s]]$$

Si  $X_t - X_s$  converge en distribución a  $X_{t-s} \sim (0, |t-s|)$

$$\mathbb{E}[X_t X_s] = \frac{1}{2}[\text{var}[X_t] + \text{var}[X_s] - \text{var}[X_{t-s}]]$$

Por condición:  $X_t = t^H X_1$

$$\mathbb{E}[X_t X_s] = \frac{1}{2}[\text{var}[t^H X_1] + \text{var}[s^H X_1] - \text{var}[|t-s|^H X_1]]$$

$$\mathbb{E}[X_t X_s] = \frac{1}{2}[t^{2H} \text{var}[X_1] + s^{2H} \text{var}[X_1] - |t-s|^{2H} \text{var}[X_1]]$$

$$\mathbb{E}[X_t X_s] = \frac{1}{2}[t^{2H} + s^{2H} - |t-s|^{2H}] \text{var}[X_1]$$

$$\text{cov}(X_t X_s) = \mathbb{E}[X_t X_s] - \mathbb{E}[X_t] \mathbb{E}[X_s]$$

Debido a que son procesos gaussianos centrados

$$\text{cov}(X_t X_s) = \mathbb{E}[X_t X_s]$$

Como conclusión el proceso  $\{X_t\}_{t \in T}$  autosimilar con  $0 < H < 1$  con incrementos estacionarios y varianza finita tiene la misma covarianza que el movimiento Browniano fraccional con excepción de  $\text{var}(X_1)$  pero se considera movimiento Browniano fraccional.



3. Si  $H = 1$  entonces  $\mathbb{E}[X_t - tX_1]^2 = 0$

Entonces si  $\text{var}[X_t - tX_1] = \mathbb{E}[(X_t - tX_1)^2]$

$$\begin{aligned} & \mathbb{E}[X_t]^2 - 2\mathbb{E}[X_t tX_1] + \mathbb{E}[tX_1]^2 \\ & \mathbb{E}[t^H X_1]^2 - 2\mathbb{E}[t^H X_1 tX_1] + t^2 \mathbb{E}[X_1^H]^2 \\ & \mathbb{E}[t^H X_1]^2 - 2\mathbb{E}[t^H X_1 tX_1] + t^2 \mathbb{E}[X_1^H]^2 \end{aligned}$$

si  $H=1$ , entonces

$$\begin{aligned} & \mathbb{E}[tX_1]^2 - 2\mathbb{E}[tX_1 tX_1] + t^2 \mathbb{E}[X_1^H]^2 \\ & t^2 \mathbb{E}[X_1]^2 - 2\mathbb{E}[t^2 X_1^2] + t^2 \mathbb{E}[X_1^H]^2 \\ & t^2 \mathbb{E}[X_1]^2 - 2t^2 \mathbb{E}[X_1^2] + t^2 \mathbb{E}[X_1^H]^2 \\ & [t^2 - 2t^2 + t^2] \mathbb{E}[X_1]^2 \end{aligned}$$

Por lo tanto:

$$\mathbb{E}[X_t - tX_1]^2 = 0$$

**Proposición 2.3.18** *Un proceso autosimilar con índice de autosimilaridad  $H \in (0, 1)$  con incrementos estacionarios es un MBF. Si  $\text{var}(X_1) = 1$  entonces dicho proceso se denomina Movimiento Browniano fraccionario estándar*

**Proposición 2.3.19** *Un proceso  $\{X_t\}_{t \in T}$  autosimilar y gaussiano con incrementos estacionarios con  $\text{var}(X_t) = t^{2H}$  con  $H > 0$  y  $X_0 = 0$  entonces se denomina movimiento Browniano fraccionario*

Entonces se probará que la covarianza del proceso gaussiano coincide con la covarianza del movimiento Browniano fraccional, se consideraran :

1.  $X_0 = 0$
2. Incrementos estacionarios
3. proceso gaussiano centrado

$$\text{var}(X_t - X_s) = \mathbb{E}[X_t - X_s]^2$$

$$\mathbb{E}[X_t^2 - 2X_t X_s + X_s^2]$$

$$\mathbb{E}[X_t^2] - 2\mathbb{E}[X_t X_s] + \mathbb{E}[X_s^2]$$

$$\text{var}(X_t - X_s) = \mathbb{E}[X_t^2] - 2\mathbb{E}[X_t X_s] + \mathbb{E}[X_s^2]$$

$$2\mathbb{E}[X_t X_s] = \mathbb{E}[X_t^2] + \mathbb{E}[X_s^2] - \text{var}(X_t - X_s)$$

$$\mathbb{E}[X_t X_s] = \frac{1}{2}[\mathbb{E}[X_t^2] + \mathbb{E}[X_s^2] - \text{var}(X_t - X_s)]$$

$$\mathbb{E}[X_t X_s] = \frac{1}{2}[\text{var}[X_t] + \text{var}[X_s] - \text{var}(X_t - X_s)]$$

Al reemplazar con:

$$\mathbb{E}[X_t X_s] = \frac{1}{2}[t^{2H} + s^{2H} - |t - s|^{2H}]$$

$\forall s, t \in T$

Por lo tanto la covarianza del proceso  $\{X_t\}_{t \in T}$  coincide con la covarianza del movimiento Browniano fraccional entonces el proceso  $\{X_t\}_{t \in T}$  es un movimiento Browniano fraccional con índice de autosimilaridad  $H$ .

**Corolario 2.3.20** *Sea  $\{X_t\}_{t \in T}$  un proceso con varianza finita con  $H \in (0, 1]$  entonces:*

*$\{x_t\}_{t \in T}$  un proceso autosimilar gaussiano con índice  $H$  y con incrementos estacionarios.*

*$\{x_t\}_{t \in T}$  es un movimiento Browniano fraccionario con índice de autosimilaridad  $H$*

*$\{x_t\}_{t \in T}$  es gaussiano centrado si  $H \in (0, 1)$  con operador estadístico de covarianza*

$$\text{cov}(X_t, X_s) = \frac{1}{2}[t^{2H} + s^{2H} - |t - s|^{2H}]\text{var}[X_1], \forall t, s \in T \quad (2.5)$$

## $\alpha$ - variación del movimiento Browniano fraccionario

Sea  $\{X_t\}_{t \in [0, T]}$  es un proceso estocástico donde  $T$  es una constante entonces tenemos al intervalo de tiempo  $[0, T]$ , entonces:

**Proposición 2.3.21** *Sea un proceso estocástico  $\{X_t\}_{t \in [0, T]}$  decimos que tiene  $\alpha$  - variación finita, si la sucesión de variables aleatorias*

$$S_\alpha(X, \pi_n) = \sum_{i=1}^n |X_{t_i} - X_{t_{i-1}}|^\alpha$$

,  $n \in \mathbb{N}$

*Para toda familia de particiones  $\pi_n = \{0 = t_0 < t_1 < \dots < t_n = T\}$  converge en probabilidad a una variable aleatoria  $V_\alpha(X, [0, T])$  entonces a dicha variable aleatoria se le denomina como  $\alpha$ -variación del proceso  $\{X_t\}_{t \in [0, T]}$*

Entonces para calcular la  $\alpha$ -variación del movimiento Browniano fraccional sobre el intervalo  $[0, T]$ , se utiliza una sucesión de particiones  $\pi_n$ , en donde:

$$\pi_n = \left\{ \frac{iT}{n}, i = 0, 1, 2, \dots, n \right\}$$

,  $n \in \mathbb{N}$

Para calcular la sucesión de variables aleatorias

$$S_\alpha(B^H, \pi_n) = \sum_{i=1}^n |B_{\frac{iT}{n}} - B_{\frac{(i-1)T}{n}}|^\alpha$$

Porque segun la proposición dicha suma converge en probabilidad a  $V_\alpha(X, [0, T])$  la cual es denominada  $\alpha$  - variación del proceso  $\{X_t\}_{t \in [0, T]}$  donde  $T$  es una constante.

Entonces se tendra que desarrollar la sucesión de variables aleatorias

$$S_\alpha(X, \pi_n) = \sum_{i=1}^n |X_{t_i} - X_{t_{i-1}}|^\alpha$$

utilizando el principio de estacionariedad y el principio de autosimilitud, tenemos

$$S_\alpha(B^H, \pi_n) = \sum_{i=1}^n |B_{\frac{iT}{n}}^H - B_{\frac{(i-1)T}{n}}^H|^\alpha$$

Por el principio de estacionariedad

$$S_\alpha(B^H, \pi_n) = \sum_{i=1}^n |B_{\frac{T}{n}}^H|^\alpha$$

Por el principio de autosimilitud

$$S_\alpha(B^H, \pi_n) = \sum_{i=1}^n \left| \left( \frac{1}{n} \right)^H B_T^H \right|^\alpha$$

$$\begin{aligned}
S_\alpha(B^H, \pi_n) &= \left(\frac{1}{n}\right)^{H\alpha} \sum_{i=1}^n |B_T^H|^\alpha \\
S_\alpha(B^H, \pi_n) &= \left(\frac{1}{n}\right)^{H\alpha} n |B_T^H|^\alpha \\
S_\alpha(B^H, \pi_n) &= \left(\frac{1}{n}\right)^{(H\alpha-1)} |B_T^H|^\alpha
\end{aligned}$$

Antes de calcular la variación del proceso del movimiento Browniano fraccionario se analizará:

$$S_\alpha(B^H, \pi_n) \rightarrow_{n \rightarrow \infty} \begin{cases} 0 & \text{si } H\alpha > 1 \\ \infty & \text{si } H\alpha < 1 \\ |B_T^H|^\alpha & \text{si } H\alpha = 1 \end{cases}$$

Si calculamos la  $\alpha$  - variación del proceso del movimiento Browniano fraccionario, entonces

$$V_\alpha(X, [0, T]) = \left(\frac{1}{n}\right)^{(H\alpha-1)} |B_T^H|^\alpha$$

$B_T^H = T^H B_1^H$ , entonces

$$V_\alpha(X, [0, T]) = \left(\frac{1}{n}\right)^{(H\alpha-1)} |T^H B_1^H|^\alpha$$

$$V_\alpha(X, [0, T]) = \left(\frac{1}{n}\right)^{(H\alpha-1)} T^{H\alpha} |B_1^H|^\alpha$$

Entonces la  $\alpha$ -variación del movimiento Browniano fraccional sera:

$$V_\alpha(X, [0, T]) \rightarrow_{n \rightarrow \infty} \begin{cases} 0 & \text{si } H\alpha > 1 \\ \infty & \text{si } H\alpha < 1 \\ T |B_1^H|^\alpha & \text{si } H\alpha = 1 \end{cases} \quad (2.6)$$

**Proposición 2.3.22** *La variable aleatoria  $T\mathbb{E}(|N|^\alpha)$  con  $N \sim N(0, 1)$  donde  $\alpha = \frac{1}{H}$  es la  $\alpha$ -variación del movimiento Browniano fraccional en el conjunto de tiempo  $[0, T]$ . (Demostracion Tudor, 2002)*

Se concluye que las trayectorias del proceso tienen  $\alpha$ -variación no acotadas para  $H\alpha < 1$  y para  $H > \frac{1}{2}$  su variación cuadrática es cero (Nualart 2003), derivado de la expresión anterior se concluye que el movimiento Browniano fraccionario para  $H \neq \frac{1}{2}$  no es una semimartingala (Lin 1995).

Entonces si  $V_\alpha(X, [0, T]) = \left(\frac{1}{n}\right)^{(H\alpha-1)} T^{H\alpha} |B_1^H|^\alpha$  con  $\alpha = 2$  se tiene:

$$V_2(X, [0, T]) = \left(\frac{1}{n}\right)^{(2H-1)} T^{2H} |B_1^H|^\alpha$$

Por lo tanto:

$$V_2(X, [0, T]) \rightarrow_{n \rightarrow \infty} \begin{cases} 0 & \text{si } H > \frac{1}{2} \\ \infty & \text{si } H < \frac{1}{2} \\ T|B_T^H|^2 & \text{si } H = \frac{1}{2} \end{cases} \quad (2.7)$$

Si en un proceso de Markov la probabilidad del futuro solo depende del presente por lo cual estos procesos se les llama procesos sin memoria o con pérdida de memoria.

**Teorema 2.3.23** *El movimiento Browniano fraccional con índice de similitud  $H \in (0, 1)$ , no es un proceso de Markov para  $H \neq \frac{1}{2}$  (Demostración en Tudor, 2002)*

## Diferenciabilidad de las trayectorias

**Proposición 2.3.24** . *Casi todas las trayectorias del movimiento Browniano Fraccionario no son diferenciables en ningún punto. (Azucena Campillo, 2013).*

*Demostración.* Desde la Definición de Paul Levy, si consideramos que análogo a la singularidad un proceso estocástico con tiempo  $t$  fijo entonces el proceso es una variable aleatoria en un espacio de probabilidad. Por lo tanto suponga que  $B_0^H = 0$  y sea la variable aleatoria :

$$\frac{B_t^H - B_{t_0}^H}{(t - t_0)}$$

, con  $\forall t > t_0$

sabiendo que el movimiento browniano fraccional es un proceso autosimilar con incrementos estacionarios entonces se cumple que:

$$\frac{B_t^H - B_{t_0}^H}{t - t_0} \equiv \frac{B_{t-t_0}^H}{t - t_0} \equiv \frac{(t - t_0)^H B_1^H}{(t - t_0)} \equiv (t - t_0)^{H-1} B_1^H$$

Si consideramos  $w$  fijo entonces el proceso es una  $W$ - trayectoria una función, por lo tanto consideramos el evento  $A_t$ .

$$A_t = \{w : \sup_{t_0 \leq s \leq t} \left| \frac{B_s^H(w) - B_{t_0}^H(w)}{(s - t_0)} \right| > d\}$$

$$\forall d \in \mathbb{R}_+$$

Entonces para cualquier sucesión decreciente  $(t_n)_{n \in \mathbb{N}}$  a  $t_0$ , entonces :

$$A_t = \lim_{n \rightarrow \infty} A_{t_n} = \cap_{n=1}^{\infty} A_{t_n}$$

$$A_{t_n} \supset A_{t_{n+1}}$$

$$A_{t_n} \supset \{w : \left| \frac{B_{t_n}^H(w) - B_{t_0}^H(w)}{(t_n - t_0)} \right| > d\}$$

$$A_{t_n} = \{|(t_n - t_0)^{H-1} B_1^H(w)| > d\}$$

$$A_{t_n} = \{|B_1^H(w)| > (t_n - t_0)^{1-H} d\}$$

Por lo tanto si le aplicamos medida

$$\mathbb{P}(A_{t_n}) \geq \mathbb{P}(|B_1^H| > (t_n - t_0)^{1-H} d)$$

Si tomamos límites , entonces tenemos:

$$\begin{aligned}\lim_{n \rightarrow \infty} \mathbb{P}(A_{t_n}) &\geq \lim_{n \rightarrow \infty} \mathbb{P}(|B_1^H| > (t_n - t_0)^{1-H} d) \\ &= \mathbb{P}\{\cup_{n=1}^{\infty} (|B_1^H| > (t_n - t_0)^{1-H} d)\}\end{aligned}$$

con  $H \in (0, 1)$

$$\lim_{n \rightarrow \infty} \mathbb{P}(A_{t_n}) = \mathbb{P}(|B_1^H| > 0) = 1$$

Por la continuidad secuencial en probabilidad

$$\lim_{n \rightarrow \infty} \mathbb{P}(A_{t_n}) = \mathbb{P}(\lim_{n \rightarrow \infty} A_{t_n})$$

Por lo tanto

$$\mathbb{P}(\lim_{n \rightarrow \infty} A_{t_n}) = \mathbb{P}(\lim_{n \rightarrow \infty} w : \sup_{t_0 \leq s \leq t} \left| \frac{B_s^H(w) - B_{t_0}^H(w)}{(s - t_0)} \right| > d) = 1$$

$$\mathbb{P}(\lim_{t_n \rightarrow t_0} A_{t_n}) = 1$$

Esto implica que en el evento en donde  $B^H$  no es diferenciable en el punto  $t_0 \in T$  tiene medida 1, con  $t_0$  arbitrario. Entonces se comprueba la proposición de no diferenciabilidad de las trayectorias del movimiento browniano fraccional en ningún punto.

## Dependencia de largo plazo

Se considera que una sucesión de variables aleatorias  $\{\omega_n\}_{n \in \mathbb{N}}$  tiene dependencia de largo plazo si sus covarianzas  $\rho(n) = \text{cov}(\omega_k, \omega_{k+n})$  satisfacen:

$$\lim_{n \rightarrow \infty} \frac{\rho(n)}{cn^{-\alpha}} = 1$$

,  $\forall k, n \in \mathbb{N}$ , donde  $c$  es una constante y  $\alpha \in (0, 1)$

Entonces se considera que:

$\rho(n) \sim cn^{-\alpha}$ , cuando  $n \rightarrow \infty$ , donde  $n^{-\alpha}$  es la tasa a la que decrece la autocorrelación de manera hiperbólica, es decir lentamente por lo que  $\sum_{k=0}^{\infty} \rho(n)$  no es absolutamente sumable.

Si incluimos el parámetro que me indica el grado de memoria  $d$ , tenemos:

$$\rho(n) \sim cn^{2d-1}, \text{ cuando } n \rightarrow \infty, \forall d \in (0, 1)$$

Entonces denotamos:  $2d - 1 = -\alpha \Rightarrow d = \frac{1-\alpha}{2}$ , se observa una relación inversa entre el nivel de persistencia o memoria de un proceso y el valor de  $\alpha$ ,  $d = 0,5 - 0,5\alpha$ , entonces para  $\alpha \in (0, 1)$ , tenemos:

$\alpha = 1 \rightarrow d = 0 \rightarrow$  Movimiento Browniano (GARCH), (incrementos independientes)

A mayor grado de memoria del proceso menor valor de  $\alpha$ .

Para cada  $k, n \in \mathbb{N}$  sean:

$$\begin{aligned}\omega_k &= B_k^H - B_{k-1}^H \\ \omega_{k+n} &= B_{k+n}^H - B_{k+n-1}^H\end{aligned}$$

tenemos el vector aleatorio,  $(B_k^H - B_{k-1}^H, B_{k+n}^H - B_{k+n-1}^H)$

Para probar que el proceso Estocástico  $\{\omega_n\}_{n \in \mathbb{N}}$  tiene dependencia de largo plazo, si  $H > \frac{1}{2}$ , entonces se desarrolla:

$$\rho(n) = \text{cov}(\omega_k, \omega_{k+n})$$

$$\rho(n) = \text{cov}(B_k^H - B_{k-1}^H, B_{k+n}^H - B_{k+n-1}^H)$$

$$\rho(n) = \text{cov}(B_k^H, B_{k+n}^H) - \text{cov}(B_k^H, B_{k+n-1}^H) - \text{cov}(B_{k-1}^H, B_{k+n}^H) + \text{cov}(B_{k-1}^H, B_{k+n-1}^H)$$

$$K_H(k, k+n) - K_H(k, k+n-1) - K_H(k-1, k+n) + K_H(k-1, k+n-1)$$



$$\begin{aligned}
& \frac{1}{2} [|k|^{2H} + |k+n|^{2H} - |n|^{2H}] - \frac{1}{2} [|k|^{2H} + |k+n-1|^{2H} - |n-1|^{2H}] - \frac{1}{2} [|k-1|^{2H} + |k+n-1|^{2H} - |n|^{2H}] \\
& \frac{1}{2} [|n-1|^{2H} + |n+1|^{2H} - 2|n|^{2H}] \\
& \frac{1}{2} [n^{2H} (1 - \frac{1}{n})^{2H} + n^{2H} (1 + \frac{1}{n})^{2H} - 2|n|^{2H}] \\
& \frac{1}{2} \{ n^{2H} [(1 - \frac{1}{n})^{2H} + (1 + \frac{1}{n})^{2H} - 2] \} \\
& \frac{n^{2H} n^{-2}}{2n^{-2}} \{ [(1 - \frac{1}{n})^{2H} + (1 + \frac{1}{n})^{2H} - 2] \} \\
& \frac{n^{2H-2}}{2} [ \frac{(1 - \frac{1}{n})^{2H} + (1 + \frac{1}{n})^{2H} - 2}{n^{-2}} ] \\
& \rho(n) = \frac{n^{2H-2}}{2} \lambda
\end{aligned}$$

Aplicamos L' Hóspital a  $\lambda$ :

$$\frac{(1 - \frac{1}{n})^{2H} + (1 + \frac{1}{n})^{2H} - 2}{n^{-2}} \rightarrow 2H(2H - 1), \text{ cuando } n \rightarrow \infty$$

Entonces:

$$\rho(n) = \frac{n^{2H-2}}{2} [2H(2H - 1)]$$

$$\rho(n) = H(2H - 1)n^{2H-2}, \forall H \in (0, 1)$$

$$\rho(n) \sim cn^{-\alpha} \text{ con } c \neq 0, \text{ entonces } c = H(2H - 1), \text{ por lo tanto } -\alpha = 2H - 2 \\ \Rightarrow H = \frac{2-\alpha}{2}$$

Si  $2d - 1 = 2H - 2 \Rightarrow d = H - 0,5$ , por lo tanto  $d$  tiene una relación directa con  $H$ , es decir a mayor nivel de memoria mayor valor de  $H$ , entonces la covarianza de los incrementos tienden a cero en el orden  $n^{2H-2}$ , entonces el movimiento Browniano fraccional es un proceso de memoria larga.

$$\lim_{n \rightarrow \infty} \frac{\rho(n)}{H(2H - 1)n^{2H-2}} = 1$$

Si  $H \in (\frac{1}{2}, 1) \Rightarrow \alpha \in (0, 1) \Rightarrow \text{MBF (dependencia de largo plazo)}$

$H \in (0, \frac{1}{2}) \Rightarrow \alpha \in (0, 1) \Rightarrow \text{MBF (no tiene dependencia de largo plazo)}$

**Corolario 2.3.25** *Un proceso estocástico se dice:*

1. *Persistente  $\Rightarrow$  sus trayectorias tienen la misma dirección, es decir:*

*$H \in (\frac{1}{2}, 1) \Rightarrow d \in (0, 0.5) \Rightarrow H(2H - 1)n^{2H-2} > 0$ ; los incrementos del MBF tienden a la misma dirección por lo tanto es un proceso persistente. existe dependencia a largo plazo*

2. *Antipersistente  $\Rightarrow$  sus trayectorias tienden a retroceder a misma, es decir:*  
 $H \in (0, \frac{1}{2}) \Rightarrow d \in (-0.5, 0) \Rightarrow H(2H-1)n^{2H-2} < 0$ ; *los incrementos del MBF tienden a direcciones opuestas por lo tanto es un proceso antipersistente o memoria corta.*
3.  $H = \frac{1}{2} \Rightarrow d=0 \Rightarrow H(2H-1)n^{2H-2} = 0$ , *los incrementos son independientes hay neutralidad en las direcciones de sus incrementos por lo tanto es un movimiento Browniano.*

### 2.3.6. Volatilidad Dinámica

El concepto de volatilidad es muy utilizado en aplicaciones financieras, se refiere a la medida de la intensidad de los cambios del precio de un activo en términos de la desviación estándar del cambio antes mencionado, es importante mencionar que en el caso de un serie estacionaria y por lo tanto con varianza constante puede dar evidencia de oscilaciones a corto plazo que es capturado por la varianza condicional y por lo tanto modelarla para hacer predicciones a corto plazo. Pero como en la realidad la volatilidad varía en el tiempo los modelos clásico no son adecuados.

Las series financiera y económicas generalmente muestran un comportamiento sinusoidal es decir existen periodos de baja y alta volatilidad, es decir es indicación heterocedasticidad, en algunos casos se consideran más interesantes las medias y las varianzas condicionadas con el objetivo de pronóstico de corto plazo y en otros casos las medias y las varianzas no condicionadas para pronósticos a largo plazo.

Un modelo adecuado para la volatilidad tiene que modelar sus características y tener la capacidad de pronosticarla, entendiéndose como características a las siguientes:

- Aglomeración de la volatilidad. En la volatilidad se suelen presentar agrupaciones por periodo. Dicho comportamiento han sido estudiados por numerosos investigadores como Baillie(1996), Ramanathan(2012),chou(1988) y Schwert(1989), es decir innovaciones inesperadas hoy tiene influencia en el valor esperado de la volatilidad en el futuro generando un proceso muy sensible a minimos cambios en las condiciones iniciales.
- Reversión a la media. Se considera que existe un nivel normal de volatilidad a la cual toda serie tiende a retornar en fisica se define como un centro de masa es decir a periodos altos de volatilidad los periodos siguientes serán menores y a periodos bajos de volatilidad los siguientes periodos tenderán a ser más altas.
- La Volatilidad asimétrica. Se denota a sensibilidad asimétrica de la varianza condicionada ante innovaciones positivas y negativas.
- Influencia de variables exógenas.En un sistema complejo se denota la multidimensionalidad que estructura su naturaleza, los mercados financieros como es logico siguan este patron.
- Distribución de probabilidad. Generalmente la distribución de probabilidades de los retornos son leptocurticas por lo cual tiene colas pesadas .

Algunos de los modelos de volatilidad ampliamente utilizados son los modelos autoregresivos condicionalmente heterocedasticos (ARCH) desarrollados por Engle (1982), modelos autoregresivos condicionalmente heterocedásticos generalizados (GARCH) desarrollados por Bollerslev(1986) y modelos autoregresivos condicionalmente heterocedasticos fraccionalmente integrados (FIGARCH) desarrollados por Baillie, Bollerslev y Mikkelsen (1996) en donde ya estamos en el

campo de las series caóticas de memoria larga muy sensibles a mínimos cambios en las condiciones iniciales.

La dependencia del segundo orden puede ser modelada mediante una función cuadrática dinámica

## Modelos ARCH(p)

Existen diversas investigaciones acerca de estos modelos de volatilidad en este apartado se desarrollara los modelos ARCH(p) basado en la estructura de Engle de 1982.

Los procesos estocásticos de esta tesis estan definidos en el espacio de probabilidad  $(\Omega, \mathfrak{F}, \mathbb{P})$  con un conjunto de tiempo  $T = \mathbb{R}_+$ , Por lo tanto se considera como un proceso estocástico con incrementos estacionarios e independientes en donde las correlaciones con el pasado desaparecen de manera exponencial, es decir la varianza condicional se estabiliza rápidamente, entonces si consideramos un proceso estocástico  $[y_t]$  y  $\Psi_t$  un sigma de álgebra la cual es un conjunto de información para todo tiempo t, generada por las variables aleatorias  $\varepsilon_t$ , es decir  $\Psi_t = \sigma[\dots, \varepsilon_{t-2}, \varepsilon_{t-1}, \varepsilon_t]$ , las cuales son los residuos del modelo, entonces en el modelo ARCH desarrollado por Engle (1982), la varianza condicional es modelada en función del cuadrado de los residuos del periodo anterior, es decir el proceso  $\varepsilon_t$  se dice es un ARCH, donde  $\beta' = [\beta_0, \beta_1, \beta_2, \beta_3, \dots, \beta_p]$ , es el vector columna de los parámetros para modelar la media y  $x_t = [1, x_{1t}, x_{2t}, x_{3t}, \dots, x_{kt}]$  es el vector fila de variables exógenas para modelar la media,  $\alpha' = [\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_p]$  es el vector columna de parámetros para modelar la varianza condicional, Entonces si consideramos un proceso ARCH un modelo de memoria corta, por lo tanto existe auto correlación entre las perturbaciones pero esta desaparece exponencialmente velozmente a cero a medida que el tiempo tiende al infinito. Por lo tanto se define a la varianza condicionada sensible a la volatilidad externa. Entonces el modelo se define:

$$y_t | \Psi_{t-1} \sim N(\mu, h_t)$$

$$\mu_t = \beta x_t$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_p \varepsilon_{t-p}^2$$

$$\varepsilon_t = y_t - \beta x_t$$

$$\text{con } \alpha_0 > 0, \alpha_i \geq 0, i = 1, 2, \dots, p$$

Considerando los supuestos anteriormente fijados por el modelo entonces tenemos:

1.  $\varepsilon_t | \Psi_{t-1} \sim N(0, h_t)$
2.  $E(y_t | \Psi_{t-1}) = \mu = 0$
3.  $\varepsilon_t = y_t$
4.  $\varepsilon_t | \Psi_{t-1} \sim N(0, h_t)$

Entonces: Por lo tanto el modelo sería:

$$E(\varepsilon_t|\Psi_{t-1}) = 0, \text{var}(\varepsilon_t|\Psi_{t-1}) = h_t \quad (2.8)$$

con

$$h_t = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2$$

Entonces la varianza puede modelarse a través de un proceso AR(p) en las variables  $\mu_t = \varepsilon_t^2$ .

A partir de la ecuación 2.8, se definen:

$$E(\varepsilon_t) = E(E(\varepsilon_t|\Psi_{t-1})) = 0$$

$$\text{Var}(\varepsilon_t) = E(\varepsilon_t^2) = E(E(\varepsilon_t^2|\Psi_{t-1})) = E(h_t) = \sigma_t^2$$

entonces si sustituimos en la esperanza la ecuación de la varianza condicionada se expresa:

$$\text{Var}(\varepsilon_t) = \alpha_0 + \alpha_1 E(\varepsilon_{t-1}^2) + \alpha_2 E(\varepsilon_{t-2}^2) + \dots + \alpha_p E(\varepsilon_{t-p}^2)$$

$$\text{Var}(\varepsilon_t) = \alpha_0 + \alpha_1 \text{Var}(\varepsilon_{t-1}) + \alpha_2 \text{Var}(\varepsilon_{t-2}) + \dots + \alpha_p \text{Var}(\varepsilon_{t-p})$$

Si consideramos al proceso  $\varepsilon_t$ , estacionario, considerando la homocedasticidad en la serie  $\text{Var}(\varepsilon_t) = \text{Var}(\varepsilon_{t-k}), \forall k$

Entonces

$$\text{Var}(\varepsilon_t) = \alpha_0 + \alpha_1 \text{Var}(\varepsilon_t) + \alpha_2 \text{Var}(\varepsilon_t) + \dots + \alpha_p \text{Var}(\varepsilon_t)$$

$$\text{Var}(\varepsilon_t) - \alpha_1 \text{Var}(\varepsilon_t) - \alpha_2 \text{Var}(\varepsilon_t) - \dots - \alpha_p \text{Var}(\varepsilon_t) = \alpha_0$$

$$\text{Var}(\varepsilon_t)(1 - \alpha_1 - \alpha_2 - \dots - \alpha_p) = \alpha_0$$

$$\text{Var}(\varepsilon_t) = \frac{\alpha_0}{1 - \alpha_1 - \alpha_2 - \dots - \alpha_p} = \sigma_t^2$$

$$\text{Var}(\varepsilon_t) = \frac{\alpha_0}{1 - [\alpha_1 + \alpha_2 + \dots + \alpha_p]} = \sigma_t^2$$

Por lo tanto consideramos la siguiente restricción:

$$\sum_{i=1}^p \alpha_i < 1 \quad (2.9)$$

Para estimar los parámetros del modelo Arch, propuesto por Engle, que define a la variable aleatoria  $(y_t)_{t \in I}$ , se tiene la función de verosimilitud para la observación t-ésima :

$$f(y_t | \Psi_{t-1}) = \frac{1}{\sqrt{2\pi h_t}} e^{-\frac{1}{2} \frac{\varepsilon_t^2}{h_t}} \quad (2.10)$$

Ahora tomando logaritmo sin constantes:

$$l_t = -\frac{1}{2} \log h_t - \frac{1}{2} \frac{\varepsilon_t^2}{h_t}$$

Entonces para T observaciones:

$$l = \frac{1}{T} \sum_{t=1}^T l_t \quad (2.11)$$

Ahora para estimar la matriz de información necesitamos las hessianas, es decir los estimadores maximo verosímil se encuentran al resolver el sistema dado por las condiciones de primer orden:

$$\frac{\partial l_t}{\partial \alpha} = -\frac{1}{2h_t} \frac{\partial h_t}{\partial \alpha} + \frac{\varepsilon_t^2}{2h_t^2} \frac{\partial h_t}{\partial \alpha}$$

$$\frac{\partial l_t}{\partial \alpha} = \frac{h_t^{-1}}{2} \frac{\partial h_t}{\partial \alpha} \left( \frac{\varepsilon_t^2}{h_t} - 1 \right)$$

$$\frac{\partial l_t}{\partial \beta} = -\frac{1}{2h_t} \frac{\partial h_t}{\partial \beta} - \frac{1}{2} \left[ \frac{2\varepsilon_t}{h_t} \frac{\partial \varepsilon_t}{\partial \beta} - \frac{\varepsilon_t^2}{h_t^2} \frac{\partial h_t}{\partial \beta} \right]$$

$$\frac{\partial l_t}{\partial \beta} = -\frac{1}{2h_t} \frac{\partial h_t}{\partial \beta} - \frac{\varepsilon_t}{h_t} \frac{\partial \varepsilon_t}{\partial \beta} + \frac{\varepsilon_t^2}{2h_t^2} \frac{\partial h_t}{\partial \beta}$$

$$\text{Si } \frac{\partial \varepsilon_t}{\partial \beta} = -x_{tl}$$

$$\frac{\partial l_t}{\partial \beta} = -\frac{1}{2h_t} \frac{\partial h_t}{\partial \beta} + \frac{\varepsilon_t x_{tl}}{h_t} + \frac{\varepsilon_t^2}{2h_t^2} \frac{\partial h_t}{\partial \beta}$$

$$\frac{\partial l_t}{\partial \beta} = \frac{\varepsilon_t x_{tl}}{h_t} + \frac{1}{2h_t} \frac{\partial h_t}{\partial \beta} \left( \frac{\varepsilon_t^2}{h_t} - 1 \right)$$

Entonces para hallar las componentes de la matriz hessiana H:

$$\frac{\partial^2 l_t}{\partial \alpha \partial \alpha'} = \frac{h_t^{-1}}{2} \frac{\partial h_t}{\partial \alpha} \left( -\frac{\varepsilon_t^2}{h_t^2} \frac{\partial h_t}{\partial \alpha'} \right) + \left( \frac{\varepsilon_t^2}{h_t} - 1 \right) \frac{\partial}{\partial \alpha'} \left( \frac{h_t^{-1}}{2} \frac{\partial h_t}{\partial \alpha} \right)$$

$$\begin{aligned}\frac{\partial^2 l_t}{\partial \alpha \partial \alpha'} &= -\frac{h_t^{-1}}{2} \frac{\partial h_t}{\partial \alpha} \frac{\partial h_t}{\partial \alpha'} \left( \frac{\varepsilon_t^2}{h_t^2} \right) + \left( \frac{\varepsilon_t^2}{h_t} - 1 \right) \frac{\partial}{\partial \alpha'} \left( \frac{h_t^{-1}}{2} \frac{\partial h_t}{\partial \alpha} \right) \\ \frac{\partial^2 l_t}{\partial \alpha \partial \alpha'} &= -\frac{h_t^{-2}}{2} \frac{\partial h_t}{\partial \alpha} \frac{\partial h_t}{\partial \alpha'} \left( \frac{\varepsilon_t^2}{h_t} \right) + \left( \frac{\varepsilon_t^2}{h_t} - 1 \right) \frac{\partial}{\partial \alpha'} \left( \frac{h_t^{-1}}{2} \frac{\partial h_t}{\partial \alpha} \right)\end{aligned}\quad (2.12)$$

$$\frac{\partial^2 l_t}{\partial \beta \partial \beta'} = -\frac{h_t^{-2}}{2} \frac{\partial h_t}{\partial \beta} \frac{\partial h_t}{\partial \beta'} \left( \frac{\varepsilon_t^2}{h_t} \right) + \left( \frac{\varepsilon_t^2}{h_t} - 1 \right) \frac{\partial}{\partial \beta'} \left( \frac{h_t^{-1}}{2} \frac{\partial h_t}{\partial \beta} \right) - h_t^{-1} x_t' x_t - \frac{2\varepsilon_t x_t'}{h_t^2} \frac{\partial h_t}{\partial \beta'} \quad (2.13)$$

Se sabe que:

$$H_{\beta\beta} = \frac{\partial^2 l_t}{\partial \beta \partial \beta'}$$

$$H_{\alpha\alpha} = \frac{\partial^2 l_t}{\partial \alpha \partial \alpha'}$$

El espacio de parámetros  $\Theta = [\theta : \theta = (\alpha', \beta')]$

Y considerando la simetría de la matriz de información tenemos:

$$I_\theta = \begin{pmatrix} I_{\alpha\alpha} & 0 \\ 0 & I_{\beta\beta} \end{pmatrix} \quad (2.14)$$

Donde la matriz información se denota por:

$$I_\theta = -E\left(\frac{\partial^2 L_t}{\partial \theta \partial \theta'}\right) = -E[H]$$

Siendo H es la matriz hessiana y  $\theta$  es el vector de parámetros a estimar mediante el algoritmo de scoring fisher, debido a su mayor a su mejor manejo analítico y también a su estabilidad multiparámetros, es importante mencionar que al obtener la matriz de información se pueden hallar las varianzas estimadas de los parámetros si recordamos que la inversa de la matriz de información es la cota de rao-cramer.

Por lo tanto consideramos los siguientes hechos:

1. La matriz información se condiciona a la información pasada.
2.  $I_\theta = -E(H) = -E(E(H|\Psi_{t-1}))$
3.  $E\left(\frac{\varepsilon_t^2}{h_t}|\Psi_{t-1}\right) = 1$



$$4. E(\varepsilon_t|\Psi_{t-1}) = 0$$

Entonces es posible obtener los bloques de la matriz de información aplicando la esperanza a las hessianas

$$I_{\alpha\alpha} = \frac{1}{T} \sum_{t=1}^T -E[E(\frac{\partial^2 l_t}{\partial \alpha \partial \alpha'}|\Psi_{t-1})]$$

$$I_{\alpha\alpha} = \frac{1}{T} \sum_{t=1}^T E[\frac{h_t^{-2}}{2} \frac{\partial h_t}{\partial \alpha} \frac{\partial h_t}{\partial \alpha'}](2.15)$$

$$I_{\beta\beta} = \frac{1}{T} \sum_{t=1}^T -E[E(\frac{\partial l_t}{\partial \beta \partial \beta'}|\Psi_{t-1})]$$

$$I_{\beta\beta} = \frac{1}{T} \sum_{t=1}^T E[\frac{h_t^{-2}}{2} \frac{\partial h_t}{\partial \beta} \frac{\partial h_t}{\partial \beta'} + h_t^{-1} x_t' x_t](2.16)$$

Para obtener la ecuación 2.15 y la ecuación 2.16 se le aplica los hechos 3 y 4.

### Algoritmo de scoring fisher

Para estimar los parámetros  $\Theta = [\theta : \theta = (\alpha', \beta')]$  del modelo junto con sus varianzas se realiza mediante el algoritmo de Scoring Fisher, introducido por R.A.Fisher, en donde aplica la esperanza matemática a la segunda derivada de la función de logaritmo de verosimilitud, se aplica en este caso debido a su simplicidad analítica y a su estabilidad para la estimación en caso de multiparámetros, entonces:

$$\theta \leftarrow \theta + I(\theta)^{-1}S(\theta)$$

En muchos casos  $I(\theta)$ , es fácilmente calculable y  $I(\theta)$  es siempre una matriz definida positiva, esto generalmente estabiliza el algoritmo, pero se necesita muchas iteraciones como:

$$\theta \leftarrow \theta + \gamma[I(\theta) + S(\theta)^2]^{-1}S(\theta)$$

En este caso de n observaciones independientes e idénticamente distribuidas tenemos:

$$\theta \leftarrow \theta + I_1(\theta)^{-1}\frac{S(\theta)}{n}$$

Donde  $I_1(\theta)$  es la matriz de información de una observación.

Si realizamos una comparación entre los algoritmos de Scoring- Fisher y de Newton - Raphson, en el caso de multiparámetros, La aproximación para la ecuación de scoring seria:

$$S(\theta_0) - J(\theta_0)(\theta - \theta_0) = 0$$

Donde:

$$S(\theta)_r = \frac{\partial I(\theta)}{\partial \theta_r}$$
$$J(\theta)_{rs} = -\frac{\partial^2 I(\theta)}{\partial \theta_r \partial \theta_s}$$

$S(\theta)$ , es la gradiente y  $-J(\theta)$  la Hessiana de  $I(\theta)$ , entonces los pasos para la iteración pueden ser descritos así:

$$\theta \leftarrow \theta + J(\theta)^{-1}S(\theta)$$

Donde el estadístico Score S es un vector y la hessiana -J una matriz.

Ahora en relación con la estabilidad en el caso de multiparámetros, el algoritmo de Newton- Raphson dicho problema no solamente es la negatividad sino también la matriz puede ser singular es decir no necesariamente sus autovalores todos serán positivos, en este caso serán positivos y negativos, entonces no es invertible, Pero en el caso del algoritmo de Scoring - Fisher, dicha matriz siempre será definida positiva por lo tanto tendrá inversa, recordando el teorema 2.3.26. de Sylvester's:

**Teorema 2.3.26** *A s una matriz definida positiva si solo si  $\det(A_R) > 0$ , para todas la submatrices  $A_R$  de la forma  $\{a_{rs}\}_{r,s=1,\dots,R}$ .*

Entonces si reemplazamos  $J(\theta)$  por la esperanza de la matriz de información se tiene:

$$\theta \longleftarrow \theta + I(\theta)^{-1}S(\theta)$$

Donde,  $I(\theta)$ , es la matriz de información de Fisher la cual siempre es una matriz definida positiva, en el caso de multiparámetros se puede demostrar mediante la iteración:

$$\begin{aligned}\theta &\longleftarrow \theta + \gamma[I(\theta) + S(\theta)S(\theta)^T]^{-1}S(\theta) \\ \theta &\longleftarrow \theta + \gamma[I(\theta) + S(\theta)^T S(\theta)E]^{-1}S(\theta)\end{aligned}$$

Donde E es un matriz identidad

En el caso de la curva de la familia exponencial multiparámetros con densidad sería:

Donde  $\beta$ , es d dimensional:

$$\begin{aligned}f(x, \beta) &= b(x)e^{\theta\beta^T t(x) - c[\theta(\beta)]} \\ J(\beta) &= \frac{\partial^2[\theta(\beta) - \theta\beta^T t(X)]}{\partial\beta\partial\beta^T} \\ &= \frac{\partial^2\theta[[\tau[\theta(\beta) - t(X)] + [\frac{\partial\theta}{\partial\beta}]^T v[\theta(\beta)]\frac{\partial\theta}{\partial\beta}]}{\partial\beta\partial\beta^T}\end{aligned}$$

Donde la esperanza del primer término es cero, entonces:

$$I(\beta) = E[J(\theta)] = [\frac{\partial\theta}{\partial\beta}]^T v[\theta(\beta)] \frac{\partial\theta}{\partial\beta}$$

En el caso de multiparámetros se puede aproximar a:

$$\hat{\theta} \sim N_d(\theta, I(\theta)^{-1}) \quad (2.17)$$

Donde  $N_d$  es una distribución gaussiano d dimensional, entonces se aproximaría:

$$(\hat{\theta} - \theta)^T I(\theta)(\hat{\theta} - \theta) \sim (\hat{\theta} - \theta)^T I(\hat{\theta})(\hat{\theta} - \theta) \sim (\hat{\theta} - \theta)^T J(\hat{\theta})(\hat{\theta} - \theta) \sim \chi^2(d)$$

Si, Aplicamos el algoritmo Scoring Fisher al modelo ARCH(p), tenemos,  $\theta^{(\kappa)}$ , Es el vector de los valores de los parámetros en la iteración k-esima del algoritmo y  $q^{(\kappa)}$  es el vector :

$$q^{(\kappa)} = \frac{1}{T} \sum_{t=1}^T \frac{\partial l_t}{\partial \beta}$$

$I(\theta^{(\kappa)})$  , Es el valor de la matriz de información I evaluada en  $\theta^{(\kappa)}$

El algoritmo sería:

$$\theta^{(\kappa+1)} = \theta^{(\kappa)} + (I(\theta^{(\kappa)}))^{-1}q^{(\kappa)} \quad (2.18)$$

para  $\kappa \geq 0$

Si aceptamos la simetría de la matriz información podemos proponer un algoritmo iterado alternado de fisher scoring para obtener las estimaciones máximo verosímiles de los vectores de los parámetros  $\alpha$  y  $\beta$ .

Entonces dicho algoritmo puede expresarse de la siguiente forma:

$$\beta^{(\kappa+1)} = \beta^{(\kappa)} + (I_{\beta\beta}^{(\kappa)})^{-1} q_{\beta}^{(\kappa)}$$

$$\alpha^{(\kappa+1)} = \alpha^{(\kappa)} + (I_{\alpha\alpha}^{(\kappa)})^{-1} q_{\alpha}^{(\kappa)}$$

Donde

$\alpha^{(\kappa)}$  y  $\beta^{(\kappa)}$ , son los valores de los vectores de  $\beta$  y  $\alpha$  en la Iteración  $\kappa$ -ésima del algoritmo.

## Modelo GARCH (p,q)

Existen diversas investigaciones acerca de estos modelos de volatilidad en este apartado se desarrollará los modelos GARCH(p,q) basado en la estructura de Bollerslevde 1986.

Siguiendo la estructura ARCH de la sección anterior, tambien se considera como un proceso con incrementos estacionarios e independientes donde la varianza condicionada depende de la volatilidad externa y la volatilidad interna dicho aporte se le atribuye a Bollerslev quien en el año 1986 desarrollo una extensión del modelo ARCH introduciendo la parte recursiva, es decir en este caso expresado en la ecuación de la varianza condicional en donde depende tanto del cuadrado de los errores como de las varianzas condicionales retrasadas  $p$  periodos,  $z_t = [1, \varepsilon_{t-1}^2, \dots, \varepsilon_{t-q}^2, h_{t-1}, \dots, h_{t-p}]$ , vector de variables para la varianza ,  $\Psi_{t-1}$  es la información disponible hasta el tiempo  $t - 1$ .

Entonces el modelo GARCH(p,q) desarrollado por Bollerslev está dado por:

$$\varepsilon_t | \Psi_{t-1} \sim N(0, h_t)$$

$$y_t | \Psi_{t-1} \sim N(\mu, h_t)$$

$$\mu_t = \beta x_t$$

$$h_t = z_t \omega = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \gamma_j h_{t-j} \quad (2.19)$$

$$\varepsilon_t = y_t - \beta x_t$$

$$\text{con } \alpha_0 > 0, \alpha_i \geq 0, i = 1, 2, \dots, p$$

$$\gamma \geq 0, p \geq 0, q > 0$$

Donde:  $p > 0, q > 0, \alpha_0 > 0, \gamma_j \geq 0, \alpha_i \geq 0, i = 1, 2, 3, \dots, q, j = 1, 2, 3, \dots, p$  y  $\alpha(L)$  Y  $\beta(L)$  son operadores de rezago ,  $\alpha(L) = \alpha_1 L^1 + \alpha_2 L^2 + \alpha_3 L^3 + \dots + \alpha_q L^q$ ,  $\beta(L) = \beta_1 L^1 + \beta_2 L^2 + \beta_3 L^3 + \dots + \beta_p L^p$

$\alpha_i$  Se define como la sensibilidad de la varianza condicionada ante factores externos y  $\gamma_j$  se define como la sensibilidad de la varianza condicionada ante factores internos es decir su propia información histórica.

Si  $p=0$ , se tendría el proceso Arch(q). Si  $p=0$  y  $q=0$ , se tendría ruido blanco es decir un proceso estocástico con esperanza cero varianza finita e incorrelacionada serialmente un movimiento browniano visto en el apartado anterior.

Cumplandose la siguiente restricción, la cual es desarrollado en la sección de la metodología econométrica:  $\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \gamma_j < 1$  garantizando la estacionariedad del modelo.

## Estimación de parámetros

En este caso se tiene un nuevo espacio de parámetros como un sub espacio compacto de un espacio euclidiano:

$$\Theta = [\theta : \theta = (\beta', \omega')]$$

Entonces la función de máxima verosimilitud para T observaciones tiene la forma dada por la ecuación 2.11 :

$$l = \frac{1}{T} \sum_{t=1}^T l_t$$

Aplicando logaritmo tenemos:

$$l_t = -\frac{1}{2} \log h_t - \frac{1}{2} \frac{\varepsilon_t^2}{h_t}$$

Y derivando con respecto a los parámetros para modelar la media, sería similar al proceso Arch. Luego las condiciones de primer y segundo orden estan dadas por:

$$\frac{\partial l_t}{\partial \beta} = -\frac{1}{2h_t} \frac{\partial h_t}{\partial \beta} - \frac{1}{2} \left[ \frac{2\varepsilon_t}{h_t} \frac{\partial \varepsilon_t}{\partial \beta} - \frac{\varepsilon_t^2}{h_t^2} \frac{\partial h_t}{\partial \beta} \right]$$

$$\frac{\partial l_t}{\partial \beta} = -\frac{1}{2h_t} \frac{\partial h_t}{\partial \beta} - \frac{\varepsilon_t}{h_t} \frac{\partial \varepsilon_t}{\partial \beta} + \frac{\varepsilon_t^2}{2h_t^2} \frac{\partial h_t}{\partial \beta}$$

$$\text{Si } \frac{\partial \varepsilon_t}{\partial \beta} = -x_{tl}$$

$$\frac{\partial l_t}{\partial \beta} = -\frac{1}{2h_t} \frac{\partial h_t}{\partial \beta} + \frac{\varepsilon_t x_{tl}}{h_t} + \frac{\varepsilon_t^2}{2h_t^2} \frac{\partial h_t}{\partial \beta}$$

$$\frac{\partial l_t}{\partial \beta} = \frac{\varepsilon_t x_{tl}}{h_t} + \frac{1}{2h_t} \frac{\partial h_t}{\partial \beta} \left( \frac{\varepsilon_t^2}{h_t} - 1 \right)$$

Donde

$$\frac{\partial h_t}{\partial \beta} = -2 \sum_i^q \alpha_i \varepsilon_{t-i} x'_{t-i} + \sum_{j=1}^p \gamma_j \frac{\partial h_{t-j}}{\partial \beta} \quad (2.20)$$

$$\frac{\partial^2 l_t}{\partial \beta \partial \beta'} = -\frac{h_t^{-2}}{2} \frac{\partial h_t}{\partial \beta} \frac{\partial h_t}{\partial \beta'} \left( \frac{\varepsilon_t^2}{h_t} \right) + \left( \frac{\varepsilon_t^2}{h_t} - 1 \right) \frac{\partial}{\partial \beta'} \left( \frac{h_t^{-1}}{2} \frac{\partial h_t}{\partial \beta} \right) - h_t^{-1} x_{tl}' x_t - \frac{2\varepsilon_t x_{tl}'}{h_t^2} \frac{\partial h_t}{\partial \beta'}$$

En la ecuación 2.19 se incluye la parte recursiva. Ahora derivando con respecto a los parámetros para modelar la varianza condicionada, seria similar al proceso Arch, entonces las condiciones de primer y segundo orden estan dadas por:

$$\frac{\partial l_t}{\partial \omega} = -\frac{1}{2h_t} \frac{\partial h_t}{\partial \omega} + \frac{\varepsilon_t^2}{2h_t^2} \frac{\partial h_t}{\partial \omega}$$

$$\begin{aligned}
\frac{\partial l_t}{\partial \omega} &= \frac{h_t^{-1}}{2} \frac{\partial h_t}{\partial \omega} \left( \frac{\varepsilon_t^2}{h_t} - 1 \right) \\
\frac{\partial^2 l_t}{\partial \omega \partial \omega'} &= \frac{h_t^{-1}}{2} \frac{\partial h_t}{\partial \omega} \left( -\frac{\varepsilon_t^2}{h_t^2} \frac{\partial h_t}{\partial \omega'} \right) + \left( \frac{\varepsilon_t^2}{h_t} - 1 \right) \frac{\partial}{\partial \omega'} \left( \frac{h_t^{-1}}{2} \frac{\partial h_t}{\partial \omega} \right) \\
\frac{\partial^2 l_t}{\partial \omega \partial \omega'} &= -\frac{h_t^{-1}}{2} \frac{\partial h_t}{\partial \omega} \frac{\partial h_t}{\partial \omega'} \left( \frac{\varepsilon_t^2}{h_t^2} \right) + \left( \frac{\varepsilon_t^2}{h_t} - 1 \right) \frac{\partial}{\partial \omega'} \left( \frac{h_t^{-1}}{2} \frac{\partial h_t}{\partial \omega} \right) \\
\frac{\partial^2 l_t}{\partial \omega \partial \omega'} &= -\frac{h_t^{-2}}{2} \frac{\partial h_t}{\partial \omega} \frac{\partial h_t}{\partial \omega'} \left( \frac{\varepsilon_t^2}{h_t} \right) + \left( \frac{\varepsilon_t^2}{h_t} - 1 \right) \frac{\partial}{\partial \omega'} \left( \frac{h_t^{-1}}{2} \frac{\partial h_t}{\partial \omega} \right)
\end{aligned}
\tag{2.21}$$

Se sabe que:

$$H_{\omega\omega} = \frac{\partial^2 l_t}{\partial \omega \partial \omega'}$$

El espacio de parámetros  $\Theta = [\theta : \theta = (\omega', \beta')]$

Considerando la simetría de la matriz de información tenemos:

$$I_\theta = \begin{pmatrix} I_{\omega\omega} & 0 \\ 0 & I_{\beta\beta} \end{pmatrix} \tag{2.22}$$

En donde la matriz información se denota por:

$$I_\theta = -E\left(\frac{\partial^2 L_t}{\partial \theta \partial \theta'}\right) = -E[H]$$

Donde H es la matriz hessiana y  $\theta$  es el vector de parámetros a estimar mediante el algoritmo de scoring fisher, debido a su mayor a su mejor manejo analítico y también a su estabilidad multiparámetros, es importante mencionar que al obtener la matriz de información se pueden hallar las varianzas estimadas de los parámetros si recordamos que la inversa de la matriz de información es la cota de rao-cramer.

Debido a la simplicidad de su mejor manejo analítico y también a su estabilidad multiparámetros se utilizará la esperanza condicionada a la información pasada como en el caso de Arch.

Por lo tanto consideramos los siguientes hechos:

1. la matriz información se condiciona a la información pasada.
2.  $I_\theta = -E(H) = -E(E(H|\Psi_{t-1}))$
3.  $E(\frac{\varepsilon_t^2}{h_t}|\Psi_{t-1}) = 1$
4.  $E(\varepsilon_t|\Psi_{t-1}) = 0$

Entonces es posible obtener los bloques de la matriz de información aplicando la esperanza a las hessianas

$$I_{\beta\beta} = \frac{1}{T} \sum_{t=1}^T -E[E(\frac{\partial l_t}{\partial \beta \partial \beta'}|\Psi_{t-1})]$$

$$I_{\beta\beta} = \frac{1}{T} \sum_{t=1}^T E[\frac{h_t^{-2}}{2} \frac{\partial h_t}{\partial \beta} \frac{\partial h_t}{\partial \beta'} + h_t^{-1} x_t' x_t]$$

La cual será estimada consistentemente por su primera derivada, entonces tenemos:

$$\hat{I}_{\beta\beta} = \frac{1}{T} \sum_{t=1}^T [\frac{h_t^{-2}}{2} \frac{\partial h_t}{\partial \beta} \frac{\partial h_t}{\partial \beta'} + h_t^{-1} x_t' x_t] \quad (2.23)$$

Luego como en el proceso Arch en Garch se estima la matriz información para los parametros de la varianza:

$$I_{\omega\omega} = \frac{1}{T} \sum_{t=1}^T -E[E(\frac{\partial l_t}{\partial \omega \partial \omega'}|\Psi_{t-1})]$$

$$I_{\omega\omega} = \frac{1}{T} \sum_{t=1}^T E(\frac{h_t^{-2}}{2} \frac{\partial h_t}{\partial \omega} \frac{\partial h_t}{\partial \omega'})$$

La cual es estimada consistentemente por su análogo muestral que involucra unicamente las primeras derivadas:

$$\hat{I}_{\omega\omega} = \frac{1}{T} \sum_{t=1}^T (\frac{h_t^{-2}}{2} \frac{\partial h_t}{\partial \omega} \frac{\partial h_t}{\partial \omega'}) \quad (2.24)$$



## Algoritmo de scoring fisher

Para estimar los parámetros en el espacio de parámetros  $\Theta = [\theta : \theta = (\omega, \beta)]$  del modelo junto con sus varianzas se realiza mediante el algoritmo de Scoring Fisher, similar a la estimación del modelo ARCH de la sección anterior, se aplica la esperanza matemática a la segunda derivada de la función de logaritmo de verosimilitud, se aplica en este caso debido a su simplicidad analítica y a su estabilidad para la estimación en caso de multiparámetros, entonces:

$$\theta \leftarrow \theta + I(\theta)^{-1}S(\theta)$$

En muchos casos  $I(\theta)$ , es fácilmente calculable y  $I(\theta)$  es siempre una matriz definida positiva, esto generalmente estabiliza el algoritmo, pero se necesita muchas iteraciones como:

$$\theta \leftarrow \theta + \gamma[I(\theta) + S(\theta)^2]^{-1}S(\theta)$$

En este caso de  $n$  observaciones independientes e idénticamente distribuidas tenemos:

$$\theta \leftarrow \theta + I_1(\theta)^{-1} \frac{S(\theta)}{n}$$

Donde  $I_1(\theta)$  es la matriz de información de una observación.

El análisis es similar al modelo ARCH para la comparación entre los algoritmos de Scoring- Fisher y de Newton - Raphson, en el caso de multiparámetros, La aproximación para la ecuación de scoring sería:

$$S(\theta_0) - J(\theta_0)(\theta - \theta_0) = 0$$

Donde:

$$S(\theta)_r = \frac{\partial I(\theta)}{\partial \theta_r}$$
$$J(\theta)_{rs} = -\frac{\partial^2 I(\theta)}{\partial \theta_r \partial \theta_s}$$

$S(\theta)$ , es la gradiente y  $-J(\theta)$  la Hessiana de  $I(\theta)$ , entonces los pasos para la iteración pueden ser descritos así:

$$\theta \leftarrow \theta + J(\theta)^{-1}S(\theta)$$

El estadístico Score  $S$  es un vector y la hessiana  $-J$  una matriz.

Ahora en relación con la estabilidad en el caso de multiparámetros, el algoritmo de Newton- Raphson dicho problema no solamente es la negatividad sino también la matriz puede ser singular es decir no necesariamente sus autovalores todos serán positivos, en este caso serán positivos y negativos, entonces no es invertible, Pero en el caso del algoritmo de Scoring - Fisher, dicha matriz siempre será definida positiva por lo tanto tendrá inversa, por el teorema 2.3.26.

Entonces si reemplazamos  $J(\theta)$  por la esperanza de la matriz de información se tiene:

$$\theta \longleftarrow \theta + I(\theta)^{-1}S(\theta)$$

Donde,  $I(\theta)$ , es la matriz de información de Fisher la cual siempre es una matriz definida positiva, en el caso de multiparámetros se puede demostrar mediante la iteración:

$$\begin{aligned}\theta &\longleftarrow \theta + \gamma[I(\theta) + S(\theta)S(\theta)^T]^{-1}S(\theta) \\ \theta &\longleftarrow \theta + \gamma[I(\theta) + S(\theta)^TS(\theta)E]^{-1}S(\theta)\end{aligned}$$

Donde E es un matriz identidad

En el caso de la curva de la familia exponencial multiparámetros con densidad seria:

Donde  $\beta$ , es d dimensional:

$$\begin{aligned}f(x, \beta) &= b(x)e^{\theta\beta^T t(x) - c[\theta(\beta)]} \\ J(\beta) &= \frac{\partial^2[\theta(\beta) - \theta\beta^T t(X)]}{\partial\beta\partial\beta^T} \\ &= \frac{\partial^2\theta[[\tau[\theta(\beta) - t(X)] + [\frac{\partial\theta}{\partial\beta}]^T v[\theta(\beta)]\frac{\partial\theta}{\partial\beta}]}{\partial\beta\partial\beta^T}\end{aligned}$$

La esperanza del primer término es cero, entonces:

$$I(\beta) = E[J(\theta)] = [\frac{\partial\theta}{\partial\beta}]^T v[\theta(\beta)]\frac{\partial\theta}{\partial\beta}$$

En el caso de multiparámetros se puede aproximar a:

$$\hat{\theta} \sim N_d(\theta, I(\theta)^{-1}) \quad (2.25)$$

Donde  $N_d$  es una distribución gaussiano d dimensional, entonces se aproximaría:

$$(\hat{\theta} - \theta)^T I(\theta)(\hat{\theta} - \theta) \sim (\hat{\theta} - \theta)^T I(\hat{\theta})(\hat{\theta} - \theta) \sim (\hat{\theta} - \theta)^T J(\hat{\theta})(\hat{\theta} - \theta) \sim \chi^2(d)$$

Si aplicamos el algoritmo Scoring Fisher al modelo ARCH(p), tenemos,  $\theta^{(\kappa)}$ , Es el vector de los valores de los parámetros en la iteración k-esima del algoritmo y  $q^{(\kappa)}$  es el vector :

$$q^{(\kappa)} = \frac{1}{T} \sum_{t=1}^T \frac{\partial l_t}{\partial \beta}$$

$I(\theta^{(\kappa)})$  , Es el valor de la matriz de información I evaluada en  $\theta^{(\kappa)}$   
El algoritmo sería:

$$\theta^{(\kappa+1)} = \theta^{(\kappa)} + (I(\theta^{(\kappa)}))^{-1}q^{(\kappa)} \quad (2.26)$$

para  $\kappa \geq 0$

Si aceptamos la simetría de la matriz información podemos proponer un algoritmo iterado alternado de fisher scoring para obtener las estimaciones máximo verosímiles de los vectores de los parámetros  $\omega$  y  $\beta$ .

Entonces dicho algoritmo puede expresarse de la siguiente forma:

$$\beta^{(\kappa+1)} = \beta^{(\kappa)} + (I_{\beta\beta}^{(\kappa)})^{-1} q_{\beta}^{(\kappa)}$$

$$\omega^{(\kappa+1)} = \omega^{(\kappa)} + (I_{\omega\omega}^{(\kappa)})^{-1} q_{\omega}^{(\kappa)}$$

Donde

$\omega^{\kappa}$  y  $\beta^{\kappa}$ , son los valores de los vectores de  $\beta$  y  $\omega$ , en la iteración k-esima del algoritmo.

## 2.4. Análisis Estadístico para series de tiempo con memoria larga

### 2.4.1. Análisis Fractal

En esta sección se dara a conocer los procedimientos del análisis estadístico y fractal las cuales estan conformadas por las siguientes pruebas:

1. normalidad
2. distribución de colas pesada (Pareto Lévy)
3. prueba de invarianza a escala.

Y se desarrollara la metodología del cálculo del coeficiente de Hurts para identificar la existencia de memoria larga o persistencia en la serie estudiada como tambien las pruebas de verosimilitud local de Whittle, de GPH y se utilizara el coeficiente de hurts para calcular la dimensión fractal.

### 2.4.2. Análisis Estadístico

#### Prueba normalidad

Laplace fue el primero enunciar el teorema del límite central. Luego importantes matemáticos representantes de la escuela de San Petersburgo como el precursor de la matemática indeterminista A.M. Lyapunov y sus seguidores P.L. Chebishev, A.A. Markkov estudiaron sus generalizaciones. Si consideramos N variables aleatorias (I.M. Sóbol, 1983)  $\xi_1, \xi_2, \xi_3, \dots, \xi_N$  independientes e idénticamente de modo que las distribuciones de probabilidad de estas variables coincidan. Por lo tanto se cumple que:

$$M\xi_1 = M\xi_2 = M\xi_3 = \dots = M\xi_N = m$$

$$D\xi_1 = D\xi_2 = D\xi_3 = \dots = D\xi_N = b^2$$

como la suma de todas estas variables:

$$\rho_N = \xi_1 + \xi_2 + \xi_3 + \dots + \xi_N.$$

Entonces se deduce que:

$$M\rho_N = M[\xi_1 + \xi_2 + \xi_3 + \dots + \xi_N] = Nm.$$

$$D\rho_N = D[\xi_1 + \xi_2 + \xi_3 + \dots + \xi_N] = Nb^2.$$

Por lo tanto al considerar la variable aleatoria normal  $\zeta_N$  con parámetros  $a = Nm$  y  $\sigma^2 = Nb^2$ . El teorema del límite central afirma que para cualquier intervalo  $(at, bt)$  para valores grandes de N.

$$P[at < \rho_N < bt] \approx \int_{at}^{bt} p_{\zeta_N}(x)dx.$$

Entonces se afirma que la suma  $\rho_N$  con  $N$  suficientemente grande de variables aleatorias idénticas converge aproximadamente a una normal  $p_{\rho_N}(x) \approx p_{\zeta_N}(x)$ . Siendo este teorema valido para situaciones mucho más generales: es decir los sumandos  $\xi_1, \xi_2, \xi_3, \dots, \xi_N$  pueden no ser idéntico e independientes se podría considerar como un sistema estable con igual peso cada sumando.

Siendo este teorema el que explica porque las variables aleatorias aparecen con tanta frecuencia en la práctica.

Entonces a partir del Teorema del Limite Central que establece:

**Teorema 2.4.1** Sean  $x_1, x_2, x_3, \dots, x_n$ , variables aleatorias independientes e idénticamente distribuidas y extraídas de una muestra aleatoria simple con un tamaño de muestra suficientemente grande ( $n$ ), entonces:

$$Z_n = \frac{\bar{x}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \sim Normal(0, 1)$$

Donde:

$$\bar{x}_n = \sum_{i=1}^n \frac{x_i}{n}$$

El teorema del límite central establece que si tenemos variables aleatorias independientes e idénticamente distribuidas cuando la muestra de tamaño  $n$  es suficientemente grande, la distribución de la muestra es aproximadamente normal. La curva normal está centrada alrededor de la media, la cual se representa por  $\mu$ . La variación o dispersión alrededor de la media se expresa en unidades de la desviación estándar, representada por  $\sigma$ .

A modo de aplicación del teorema, se utilizara el caso discreto de una binomial:

Supongamos un juego de lotería con un solo ganador y devuelve el importe jugado a todos los números que acaban con la misma cifra que el ganador, con un solo ensayo seguiría un distribución bernoulli  $x \sim \text{bern}(p)$  o como una  $x \sim \text{bin}(1, p)$  si se realiza  $n$  sorteos entonces la la distribución binomial  $x \sim \text{bin}(n, p)$ , es decir se ha repetido un mismo experimento  $n$  veces de manera independiente y en cada sorteo la probabilidad de éxito es de  $p=0,1$  y si consideramos  $X(n)$  en número de veces que nos han devuelto el importe jugado en  $n$  veces.

Si mediante una simulación del método de monte carlo, entendiéndose como un método numérico que resuelve problemas matemáticos mediante la simulación de variables aleatorias. Entonces si aumenta el valor de  $n$ , la función de densidad de probabilidad de la variable  $X(n)$  se aproximará a la de una normal, como se muestra gráficamente:

Si dibujamos esta función de densidad de probabilidad para  $n = 3$ , obtenemos el gráfico de la figura 2-2 siguiente:

Ahora para  $n=10$ , la función de densidad de probabilidad tendrá la forma de la figura 2-3, Si  $n=100$  ensayos la forma de densidad sera la figura 2-4. Para  $n=500$ , el gráfico de la función de probabilidad tendrá la forma de la figura 2-5.

Figura 2.2: Función de densidad de probabilidad para  $n=3$

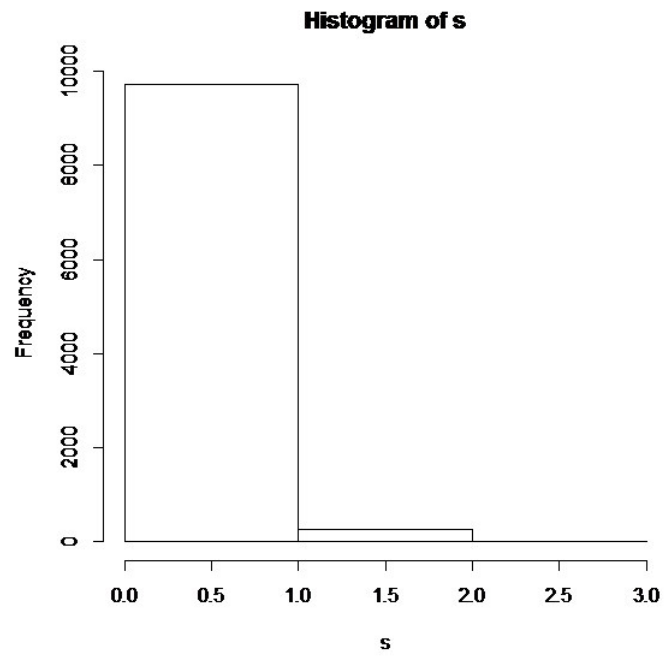


Figura 2.3: Función de densidad de probabilidad para  $n=10$

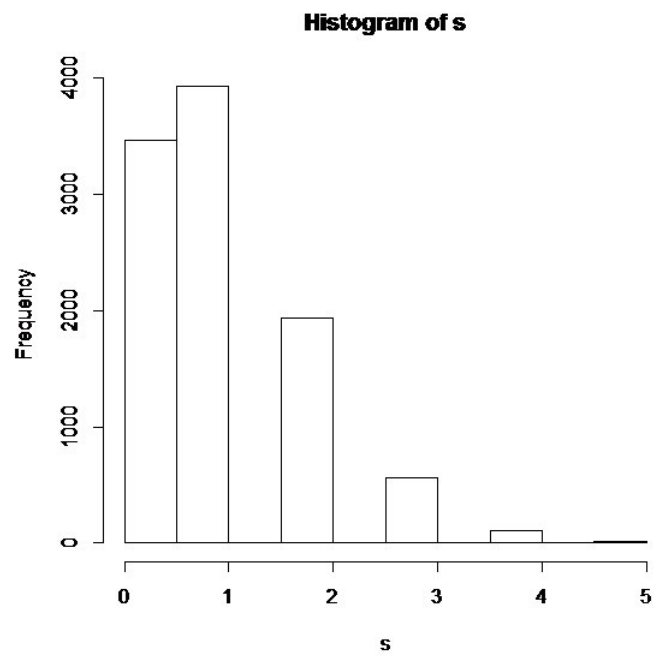


Figura 2.4: Función de densidad de probabilidad para  $n=100$

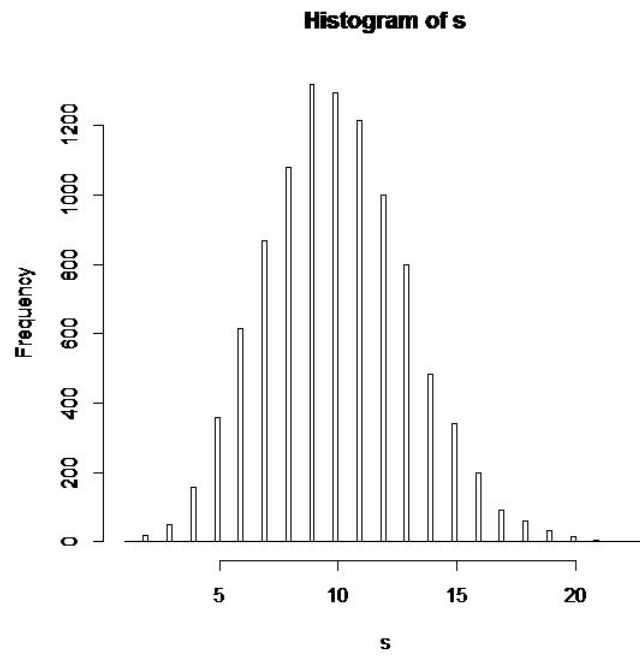
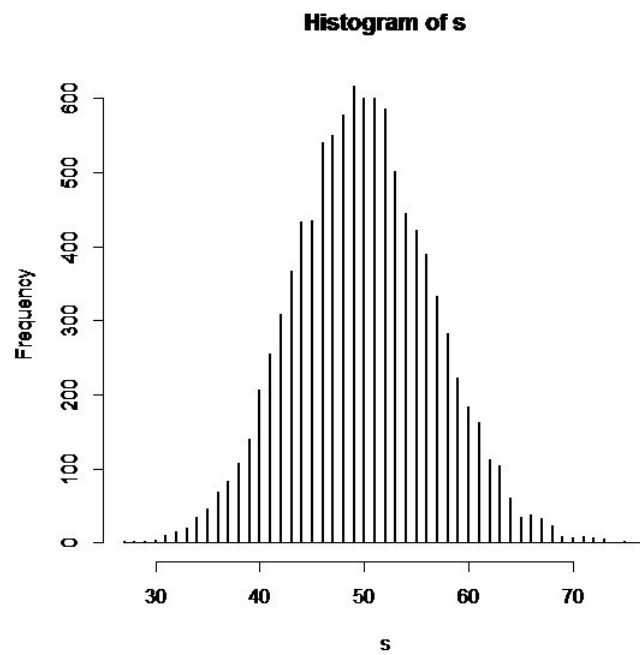


Figura 2.5: Función de densidad de probabilidad para  $n=500$



La conclusión de este experimento confirma el teorema del límite central con  $n$  suficientemente grande la variable aleatoria se aproxima a una distribución normal. con esperanza matemática de la variable

$$np = (n)(0,1)$$

y varianza

$$np(1 - p) = (n)(0,1)(0,9) = 0,09n$$

Entonces la variable aleatoria  $X(n)$  con  $n$  suficientemente grande seguirá una distribución normal con parámetros

$$X(n) \sim N(0,1n, 0,09n)$$

En finanzas, la media es su rendimiento promedio y la desviación estándar es la volatilidad. Adicionalmente a la media y a la desviación estándar, la función de distribución de probabilidad normal tiene dos características: sesgo y la curtosis, a los cuales también se les conoce como tercer y cuarto momento, respectivamente.

## Medidas de asimetría

Las medidas de asimetría o tercer momento estandar indican el grado de simetría de una distribución de valores estadísticos visualizados en el gráfico de la distribución.

1. Coeficiente de asimetría de Fisher.

$$g_1 = \frac{\frac{1}{N} \sum_{i=1}^n (x_i - \bar{x})^3 n_i}{\left( \sum_{i=1}^n (x_i - \bar{x})^2 \frac{n_i}{N} \right)^{\frac{3}{2}}}$$

Según el valor de  $g_1$ , se deduce:

- Si  $g_1 = 0$ , la distribución es simétrica.
- Si  $g_1 < 0$ , la distribución es asimétrica a la izquierda .
- Si  $g_1 > 0$ , la distribución es asimétrica a la derecha.

2. Coeficiente de asimetría de Bowley.

Está basado en los valores de los cuartiles y la mediana.

$$A_B = \frac{C_3 + C_1 - 2M_e}{C_3 - C_1}$$

Dependiendo del valor de  $A_B$  concluimos que:

- Si  $A_B = 0$ , la distribución es simétrica.
- Si  $A_B < 0$ , la distribución es asimétrica a la izquierda .
- Si  $A_B > 0$ , la distribución es asimétrica a la derecha.



3. Medida de Asimetría de Pearson.

$$A_p = \frac{\bar{x} - M_0}{S}$$

, Dado que para distribuciones campaniformes, unimodales y moderadamente asimétricas.

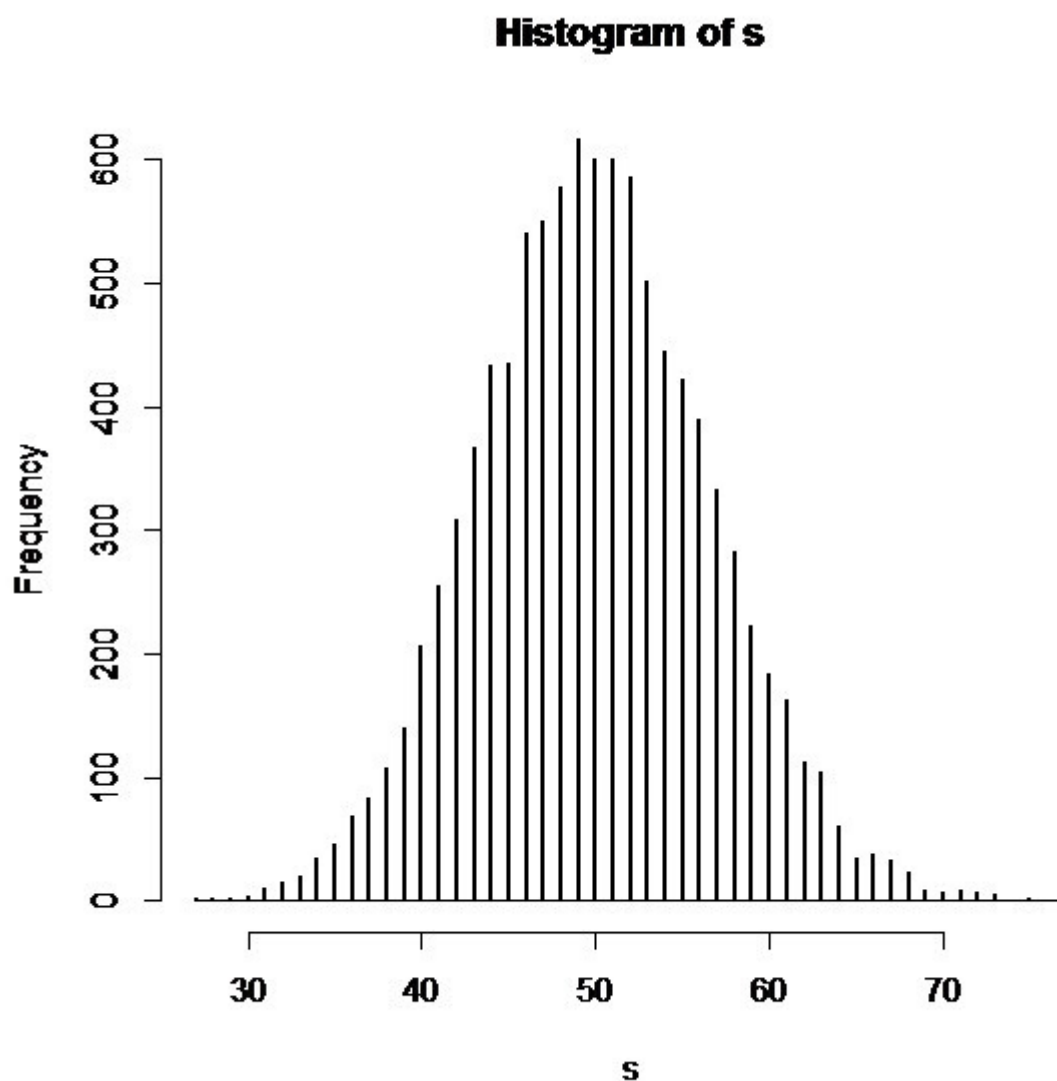
$$A_p = \frac{3(\bar{x} - M_e)}{S}$$

- Si  $A_p = 0$ , la distribución es simétrica.
- Si  $A_p < 0$ , la distribución es asimétrica a la izquierda .
- Si  $A_p > 0$ , la distribución es asimétrica a la derecha.

## Medidas de apuntamiento o curtosis

La representación gráfica de la distribución normal es:

Figura 2.6: Distribución Normal



Si tomamos como referencia la campana de Gauss se dice que otra distribución es más apuntada que la distribución normal o también con colas más pesadas (leptocúrtica) o menos apuntada con mayor dispersión mas heterogénea (platicúrtica). A las distribuciones que se asemejan a la distribución normal se les denomina mesocúrticas.

El coeficiente de apuntamiento o curtosis utilizado es el siguiente:

$$g_2 = \frac{m^4}{S^4} - 3$$

Dependiendo entonces del valor del coeficiente de  $g_2$  llamamos:

- Mesocúrtica(Normal) si , $g_2 = 0$
- Leptocurtica si , $g_2 > 0$
- Platicurtica si , $g_2 < 0$

### Prueba de distribución de Pareto-Lévy (colas pesadas)

En el análisis estadístico de la series de tiempo , es esencial para identificar si la serie obedece a una distribución gaussiana o tiene colas mas pesadas en relación a la normal por ser leptocúrtica y asimétrica(sesgada) por lo que se tiene una mejor aproximación a una distribución Pareto-Levy o una t-students o GED.

La distribución gaussiana juega un papel importante para la detección de la leptocurtosis una serie de tiempo respondiendo a una distribución de Lévy o de exceso de curtosis. Una prueba muy utilizada para dichos fines es el valor de la Curtosis, si es mayor que 3 la función es leptocúrtica, es decir se tiene gran concentración de observaciones alrededor de la media como medida de tendencia central.

### Prueba invarianza a escala.

Para este caso conceptualizamos a la invarianza de escala como una característica singular de los sistemas complejos en donde los objetos en estudio no cambia si la longitud de la escala es multiplicada por un factor constante como principio de procesos autosimilares ya estudiado en la sección de movimiento Browniano fraccional, siendo dicha característica elemento esencial de las extensiones brownianas como subfraccionario y bifraccionario singularidad del modelo FIGARCH. Si la función polinómica:  $f(x) = ax^k$ , , donde a y k son constantes, entonces  $f(cx) = c^k ax^k$ , donde c es una constante. Es decir, escalando el argumento de la función por un factor constante c, se produce un reescalamiento de la función por un factor constante  $c^k$ .

Si

$$g\left(\sum_{t=1}^n f(x)\right) = S^2$$

,

Donde:

- $S^2$ = varianza muestral de n datos
- $f(x) = ax^k$
- $x = (x_i - xm)$
- $k = 2$
- $a = \frac{1}{n}$

Por tanto, si existe invarianza de escala, la varianza re-escalada sería

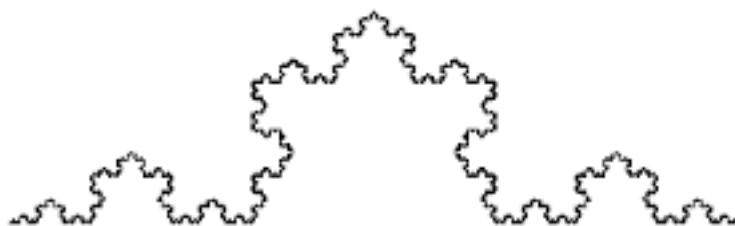
$$f(cx) = c^k ax^k$$

, por lo tanto:

$$g\left(\sum_{t=1}^n f(cx)\right) = c^2 S^2$$

Siendo  $S^2$ , la varianza de la escala original.

Figura 2.7: Autosimilitud a Escala



### 2.4.3. Pruebas de Memoria Larga

Un estadístico muy utilizado para identificar si una serie posee memoria larga es el coeficiente de Hurts también denominado coeficiente de autosimilitud, dicho instrumento fue elaborado por el hidrólogo Harold Edwing Hurst (1951), quien tenía por objetivo demostrar que el nivel de las aguas de las reservas de la represa del río Nilo no seguían un paseo aleatorio, es decir tenían memoria larga, por lo tanto persistencia lo cual indicaba una cierta dependencia no lineal entre los datos.

Hurst elaboró dicho estadístico basado en los artículos publicados en *Annalen der Physik*, n° 17, con fecha de 11 de mayo de 1905 de Albert Einstein de la teoría molecular del calor. Sobre el movimiento requerido por la teoría cinética molecular del calor de pequeñas partículas suspendidas en un líquido estacionario es decir al momento errático de una partícula de polen suspendida en un fluido de agua, Según la explicación de Einstein dichas partículas eran dinamizadas por el choque de las moléculas de agua, expresado mediante la siguiente fórmula  $R = T^{1/2}$ , donde R es el recorrido de la partícula y T el tiempo invertido en dicho recorrido, el artículo de Einstein de 1905 fue la respuesta a las observaciones del botánico Robert Brown de 1827. Es importante mencionar que tanto Albert Einstein como Robert Brown consideraban a las partículas como aleatorias e independientemente e idénticamente distribuidas fue luego del desarrollo de los matemáticos Paul Levy y Norbert Wiener que consideró para un mejor análisis ya no a las variables en si sino a la independencia de los incrementos en intervalos ajenos dando paso a los llamados procesos de Wiener la cual implica la propiedad de Markov, es decir el futuro no depende del pasado, para luego con las investigaciones de Benoit Mandelbrot y Van Ness dar lugar a los movimientos brownianos fraccionarios materia de aplicación práctica de esta tesis, afirmando que existe dependencia en los incrementos en intervalos sin tomar en cuenta la distancia entre intervalos.

#### Método del coeficiente de Hurts

El exponente de Hurts de una serie es importante porque muestra el grado de rugosidad de esta. El análisis consiste en escoger un cierto número de escalas, obtener la media del vector, obtener las diferencias de cada valor respecto a la media y hacer una suma acumulativa de ellas. El rango es la diferencia entre el valor máximo y el valor mínimo de esta suma, al dividir entre la desviación estándar se obtiene  $R/S$ , la cual según Hurst es proporcional a  $cn^H$ , donde c es una constante n tamaño de muestra y H es el coeficiente de Hurts.

Entonces el cálculo del coeficiente de Hurts viene dado por la siguiente expresión:

$$R/S(n) = an^H$$

Donde:

**a= es una constante**

**n= número de observaciones**

## H= es el exponente de Hurts

$R/S(n)$ , es el estadístico R/S , la cual esta en función de la data existente de la serie y que se define como el ratio entre el rango de variación de la serie y su desviación típica.

Dicho estadístico no paramétrico es muy utilizado para la detección de correlación con el pasado y esta dependencia es no lineal.

A continuación se describe el procedimiento del coeficiente de Hurts en los siguientes pasos (Diego Luengas Dominguez (2010), Metodología e interpretación del coeficiente de hurts):

1. Se describe la serie temporal  $X_t, t = 1, \dots, T$ .
2. Se particiona la muestra total en submuestras  $i$  de similar tamaño, donde  $n = \frac{N}{i}$ .
3. Se calcula el primer y segundo momento de la serie temporal para cada submuestra.
4. Se obtiene una nueva serie para capturar la diferencia acumulada respecto a la media del índice.
5. Se determinan el rango de cada submuestra.
6. Se obtiene el estadístico R/S (rango reescalado), mediante el siguiente cociente

$$R/S(n) = \frac{\sum_{t_0} RN, t_0}{\sum_{t_0} SN, t_0}$$

donde:

$RN, t_0$  =Rango  $SN, t_0$  =desviación típica de 3.

7. Para cada partición se calcula el promedio del rango reescalado. La potencia que determina el crecimiento de dicho rango se denomina coeficiente de Hurst, obteniéndose mediante la regresión:

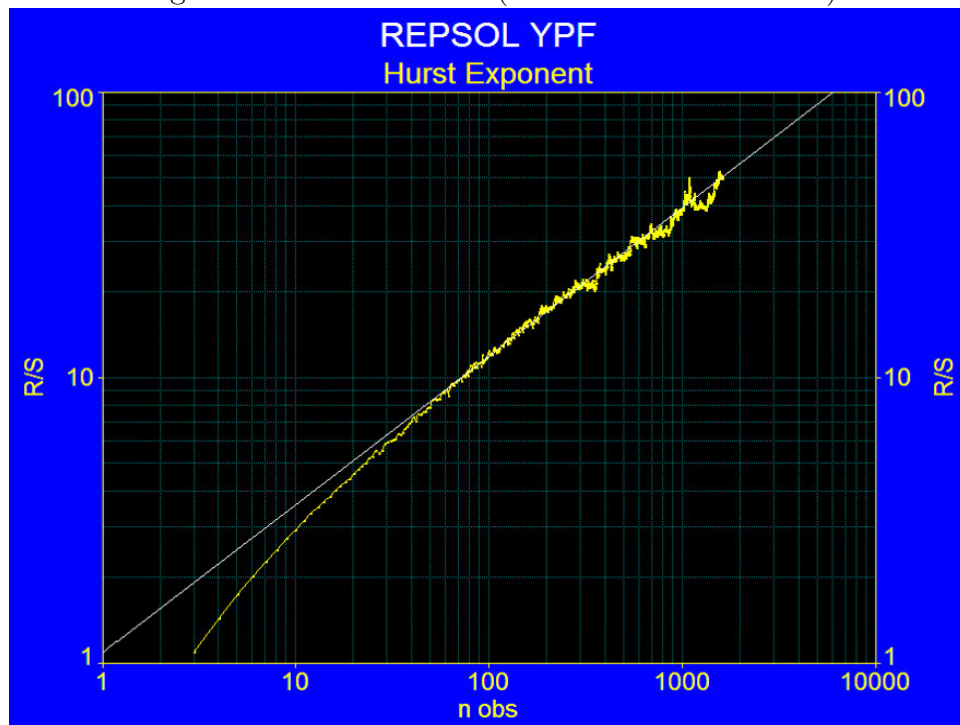
$$\log R/S = a + H \log(N)$$

## Interpretación del Exponente de Hurts

- Caso de  $H = 0,5$ , se trata de un proceso Browniano, es decir variables aleatorias independientes e idénticamente distribuidas. Se menciona entonces a un proceso gaussiano con incrementos estacionarios e independientes. Como el análisis R/S es no paramétrico por lo que pueden considerarse distribuciones otro tipo de distribuciones como la tde student, gamma, etc. Ver figura 2.8.

- En el caso de  $0,5 < H \leq 1$ , implica persistencia, es decir procesos con efectos de memoria de largo plazo. Teóricamente lo que suceda hoy impactará en el futuro por siempre se considera entonces de movimiento browniano fraccional con  $d$  como parámetro de memoria larga  $d > 0$  por lo que es un proceso con incrementos que tienden a la misma dirección. Dichos procesos son muy comunes en la naturaleza y en el entorno financiero. Por ejemplo, si  $H = 0,7$  nos indica que existe un 70 por ciento de probabilidades que el siguiente periodo siga la misma tendencia que en el periodo anterior, por otro lado una serie de tiempo persistente implica ruido negro dicho ruido genera una disminución de la de la irregularidad por lo tanto disminuirá la dimensión fractal, este ruido aparece en procesos cíclicos de largo plazo. Ver figura 2.9.
- En el caso de  $0 \leq H < 0,5$ , entonces se tiene una serie antipersistente, Un sistema antipersistente cubre menos distancia que uno aleatorio, se considera un movimiento browniano fraccional pero con parámetro de memoria  $d < 0$ . Es un proceso que necesita retroceder para volver a avanzar como un impulso por lo que los incrementos tienden a direcciones opuestas existe entonces una correlación negativa, implica ruido rosa, dicho ruido genera un incremento de la de la rugosidad por lo tanto aumentará la dimensión fractal, El ruido rosa abunda en la naturaleza y está relacionado a procesos de relajación (equilibrio dinámico) y turbulencia. Ver figura 2.10.

Figura 2.8: Ruido Blanco (Movimiento Browniano)



Dicho exponente de hurts tratado en el apartado anterior ha sido sujeto de algunas críticas, Lo(1991) muestra que la prueba estadística R/S es débil e incapaz

Figura 2.9: Ruido Negro (Persistencia)

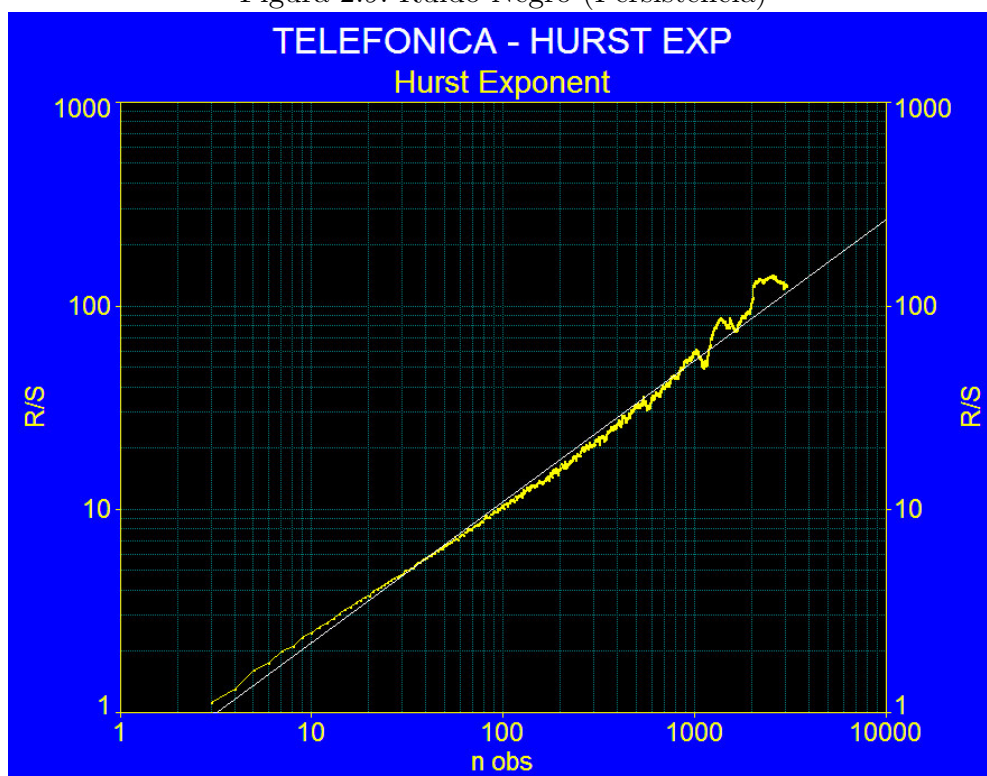


Figura 2.10: Ruido Rosa (Antipersistencia)





de distinguir si una serie tiene memoria larga o memoria corta, no encuentra evidencia que los rendimientos accionarios diarios den signos de memoria larga al modificar la prueba R/S para tomar en cuenta los efectos de la dependencia de corto plazo en los datos , lo cual no invalida del todo al exponente de Hurst en la dependencia de los datos de largo plazo.

Por tal motivo existen métodos alternativos para estimar dicho parámetro de memoria larga como se describen a continuación:

#### 2.4.4. Métodos Semi-Paramétricos

Estos métodos se basan en el supuesto de que el espectro del proceso tiene la siguiente forma(para mayor información ver César Andrés Ojeda Echeverri, 2012):

$$f(\omega) = (|1 - e^{-i\omega}|^{-2d} f^*(\omega))$$

Donde :

$f^*$ =es el componente de corto plazo en la relación de dependencia suponiéndose que es una función suave en la vecindad del origen y;

$f^*(0) = 0$ , de forma alternativa se tiene entonces:

$$f(w) = \omega^{-2d} g(\omega)$$

Donde:

$g$  es también una función suave en el origen, entonces  $g'(0) = 0$ .

Por otro lado, Geweke y Porter-Hudak (GPH)(1983)proponen un método para la estimación del parámetro de memoria larga:

$$\text{Log}\left(\sum_{j=1}^M I(\omega_{K+K-J})\right) = c + d(-2\log\sin(\frac{\omega_k}{2})) + U_k^J$$

para  $k = L + J, L + 2J, \dots, M$

Donde:

$I(\omega_j)$  =j-ésimo punto del periodograma

$$\omega_j = \frac{2\pi j}{T}$$

Considerando que  $j$  es fijo y  $L$  y  $M$  deben divergir con el tamaño de muestra, Por lo tanto el estimador que también se utilizará es conocido como el estimador local gaussiano de Whittle el cual implica en maximizar la función:

$$L_t(d) = 2d \frac{1}{M} \sum_{j=1}^M \log \omega_j - \log\left(\frac{1}{M} \sum_{j=1}^M \omega_j^{2d} I(\omega_j)\right)$$

### 2.4.5. Estimación del modelo GARCH fraccionalmente integrado (FIGARCH)

A partir del modelo desarrollado por Bollerslev en el año 1996, Baillie, Bollerslev y Mikkelsen expandieron dicho modelo para proponer un GARCH fraccionalmente integrado, con el objetivo de poder analizar la persistencia de los shocks sobre la volatilidad de las series la cual según el artículo "Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity" (Baillie, Tim Bollerslev, Hans Ole Mikkelsen, 1996), dicho efecto desaparece con el tiempo pero en forma hiperbólica es decir lentamente.

El modelo denominado FIGARCH toma en cuenta la innovación del proceso de la varianza condicionada  $v_t$ .

El modelo FIGARCH es un punto intermedio entre el modelo IGARCH en donde el efecto existe pero en este caso es permanente y el modelo GARCH, en donde el efecto existe pero desaparece exponencialmente.

Pero en FIGARCH dicho efecto su desaparición es lento ya que el orden de integración está entre  $I(0)$  y  $I(1)$ , en este modelo es común ver que existe la persistencia en la en la varianza, esto notado en la suma de los coeficientes relacionados al ARCH y GARCH los cuales se aproximan a la unidad.

La estimación es de la misma técnica que de los modelos GARCH e IGARCH, por medio del método de Quasi-MLE (véase a Andrew A. Weiss 1986).

Entonces a partir del modelo GARCH se estima el modelo FIGARCH como sigue:

$$h_t = z_t \omega = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \gamma_j h_{t-j} \quad (2.27)$$

Donde:  $p > 0, q > 0, \alpha_0 > 0, \alpha_i \geq 0, i = 1, 2, 3, \dots, q, j = 1, 2, 3, \dots, p$  y  $\alpha(L)$  Y  $\beta(L)$  son operadores de rezago,  $\alpha(L) = \alpha_1 L^1 + \alpha_2 L^2 + \alpha_3 L^3 + \dots + \alpha_q L^q$ ,  $\beta(L) = \beta_1 L^1 + \beta_2 L^2 + \beta_3 L^3 + \dots + \beta_p L^p$

podemos simplificar dicho modelo mediante la expresión:

$$h_t = \alpha_0 + \alpha(L) \varepsilon_t^2 + \beta(L) h_t \quad (2.28)$$

Si consideramos la innovación del proceso de la varianza condicional  $v_t = \varepsilon_t^2 - h_t$ , con esperanza de  $v_t$  cero,  $E[v_t] = 0$

$$E[v_t] = E[\varepsilon_t^2] + E[h_t]$$

$$E[v_t] = h_t - h_t = 0$$

$$E[v_t] = 0$$

Si despejamos la innovación de la varianza condicional, entonces tengo:

$$v_t = \varepsilon_t^2 - h_t \quad (2.29)$$

Entonces reemplazando en la ecuación de la varianza condicionada, tengo:

$$\varepsilon_t^2 - v_t = \alpha_0 + \alpha(L)\varepsilon_t^2 + \beta(L)(\varepsilon_t^2 - v_t)$$

Despejando el cuadrado de los residuos

$$\varepsilon_t^2 - v_t = \alpha_0 + \alpha(L)\varepsilon_t^2 + \beta(L)\varepsilon_t^2 - \beta(L)v_t$$

$$\varepsilon_t^2 - \alpha(L)\varepsilon_t^2 - \beta(L)\varepsilon_t^2 = \alpha_0 - \beta(L)v_t + v_t$$

$$\varepsilon_t^2(1 - \alpha(L) - \beta(L)) = \alpha_0 + v_t(1 - \beta(L))$$

Entonces este proceso puede ser expresado como un ARMA(p,q) como sigue:

$$\varepsilon_t^2(1 - \alpha(L) - \beta(L)) = \alpha_0 + v_t(1 - \beta(L))$$

Por lo tanto un proceso GARCH(p,q) integrado puede ser escrito como

$$\varepsilon_t^2(1 - \alpha(L) - \beta(L))(1 - L) = \alpha_0 + v_t(1 - \beta(L))$$

Entonces se obtiene el modelo FIGARCH reemplazando el operador  $(1 - L)$  por  $(1 - L)^d$ , el operador de diferenciación fraccional, el desarrollo completo se encuentra en la sección siguiente, donde  $0 < d < 1$

Por lo tanto el proceso gaussiano no estacionario (un proceso autosimilar no es estacionario) con incrementos dependientes cuya dependencia es no lineal con varianza definida con índice de similaridad  $H > \frac{1}{2}$  es decir es un proceso auto símil cuya volatilidad y colas pesadas es generada por esa dependencia en largas distancias de tiempo por lo tanto existe persistencia con parámetro de memoria larga  $d \in (0, 1)$  en la auto correlación del cuadrado de las innovaciones, es decir un movimiento browniano fraccional FIGARCH quedaría representado de la siguiente manera:

$$\varepsilon_t^2(1 - \alpha(L) - \beta(L))(1 - L)^d = \alpha_0 + v_t(1 - \beta(L)) \quad (2.30)$$

# Capítulo 3

## Metodología

En este capítulo se plantea el desarrollo del modelo estadístico - matemático del modelo econométrico de memoria larga denominado modelo auto regresivo con heterocedasticidad condicionada generalizado fraccionalmente integrado describiendo y analizando sus características, elementos y propiedades teóricas que permite la identificación y posterior captura de la persistencia de los shocks de la volatilidad de una serie de tiempo.

### 3.1. Diseño Metodológico

#### 3.1.1. Tipo de investigación

La presente tesis reúne las condiciones necesarias para ser denominada como una investigación estadística. Debido a la necesidad de modelar sistemas antes considerados aleatorios o movimientos brownianos geométrico supeditados a supuestos de linealidad y normalidad ahora considerados como movimiento browniano fraccionario es decir sistemas caóticos determinísticos y sus características, la presente tesis desarrolla un modelo de memoria larga con la capacidad de capturar la persistencia del efecto de la sensibilidad causada por pequeños cambios en las condiciones iniciales denominado FIGARCH.

### 3.1.2. Metodología Econométrica

### 3.1.3. Modelo autoregresivo con heterocedasticidad condicionada generalizada fraccionalmente integrado

La evidencia de la memoria larga en algunas series está asociado con el ritmo a la cual desaparece la persistencia de la auto correlación en series estacionarias, es decir a diferencia de modelos autoregresivos y de medias móviles en donde dicha función de auto correlación (ACF) desaparece exponencialmente a las cuales se la definen como series con memoria corta. En series con memoria larga, la función de auto correlación desaparece muy lentamente, generando una serie caótica, pero en ambos tipos de series, las correlaciones muestrales tienden a cero, la diferencia es la velocidad de la disminución de la covarianza de los incrementos del proceso.

Según Fan y Yao (2005), la ACF de un proceso ARMA estacionario satisface la desigualdad:

$$|\rho(k)| \leq Cr^k \quad (3.1)$$

$k=1,2,3,\dots$ , donde  $C > 0$  y  $r \in [0, 1]$ , son constantes, es decir,

$$\sum_{k=0}^{\infty} |\rho(k)| < \infty \quad (3.2)$$

por lo tanto decimos un proceso absolutamente sumable se la denomina como un proceso de memoria corta.

En cambio en los procesos de memoria larga la ACF decae muy lentamente, es decir hiperbólicamente, por tal motivo la formalidad matemática sería:

$$\rho(k) \sim Ck^{2d-1} \quad (3.3)$$

$k \rightarrow \infty$

Donde  $C \neq 0$  y  $d < 0,5$ , de esta manera la memoria larga decae muy lentamente a cero a una tasa de  $k^{2d-1}$  entonces:

$$\sum_{k=0}^{\infty} |\rho(k)| = \infty \quad (3.4)$$

la AFC no es absolutamente sumable.

Si sabemos que el proceso  $\varepsilon_t$ , denota un proceso estocástico y "t", pertenece a un conjunto de sub índices por otro lado consideramos a  $\Psi_t$ , como el conjunto de información en el tiempo "t".

Si consideramos:

$$\Psi_t = \sigma\{\dots, \varepsilon_{t-2}, \varepsilon_{t-1}, \varepsilon_t\} \quad (3.5)$$

es un sigma de álgebra sub conjunto del espacio muestral no vacío, entonces el conjunto formado por todos los subconjuntos del espacio muestral tambien es un sigma de álgebra.

Entonces se dice ser un proceso ARCH(q):

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 \quad (3.6)$$

donde  $\alpha_i \geq 0, i = 1, 2, \dots, q$

Si consideramos desde el enfoque siguiente

$$E(\varepsilon_t | \Psi_{t-1}) = 0$$

$$var(\varepsilon_t | \Psi_{t-1}) = h_t$$

$$\varepsilon_t | \Psi_{t-1} \sim N(0, h_t) \quad (3.7)$$

Por lo tanto la varianza condicional se puede expresar :

$$h_t = \sigma\{\varepsilon_{t-1}, \varepsilon_{t-2}, \varepsilon_{t-3}, \dots, \varepsilon_{t-q}, \alpha\}$$

En donde  $h(\cdot)$  es una función no negativa y  $\alpha_i = \sigma\{\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_q, \alpha\}'$  es el vector de parámetros a estimar para modelar la varianza condicionada, entonces se podría representar al proceso estocástico de la siguiente manera:

$$\varepsilon_t = h_t^{1/2} z_t, \quad (3.8)$$

en donde si desarrollamos:

$$\varepsilon_t^2 = h_t z_t^2$$

$$z_t^2 = \frac{\varepsilon_t^2}{h_t}$$

$$E[z_t^2 | \Psi_{t-1}] = E\left[\frac{\varepsilon_t^2}{h_t} | \Psi_{t-1}\right]$$

$$E[z_t^2 | \Psi_{t-1}] = 1$$

$$var(z_t | \Psi_{t-1}) = 1$$

$$E[\varepsilon_t | \Psi_{t-1}] = E[z_t | \Psi_{t-1}] E[h_t^{1/2} | \Psi_{t-1}]$$

$$0 = E[z_t | \Psi_{t-1}] E[h_t^{1/2} | \Psi_{t-1}]$$



$$E[z_t|\Psi_{t-1}] = 0$$

Entonces:

$$z_t|\Psi_{t-1} \sim N(0, 1) \quad (3.9)$$

con  $E(z_i, z_j) = 0, \forall i \neq j$

La innovación  $z_t$  tiene una distribución independiente e idénticamente distribuidas y están incorrelacionadas serialmente (movimiento browniano estándar). A partir de  $z_t$  se puede concluir que  $\varepsilon_t$  tiene una distribución con media cero y varianza  $h_t$ .

$$\varepsilon_t^2 = h_t z_t^2$$

$$E[\varepsilon_t^2|\Psi_{t-1}] = E[z_t^2|\Psi_{t-1}]E[h_t|\Psi_{t-1}]$$

$$E[\varepsilon_t^2|\Psi_{t-1}] = h_t$$

Entonces:

$$\varepsilon_t|\Psi_{t-1} \sim N(0, h_t)$$

con  $E(\varepsilon_i, \varepsilon_j) = 0, \forall i \neq j$

Si consideramos una extensión del proceso GARCH(p,q), introducido por Bollerslev en el año de 1986 y Taylor en el año de 1986, entonces tenemos:

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \gamma_j h_{t-j} \quad (3.10)$$

Donde:  $p > 0, q > 0, \alpha_0 > 0, \gamma_j \geq 0, \alpha_i \geq 0, i = 1, 2, 3, \dots, q, j = 1, 2, 3, \dots, p$  y  $\alpha(L)$  Y  $\beta(L)$  son operadores de rezago,  $\alpha(L) = \alpha_1 L^1 + \alpha_2 L^2 + \alpha_3 L^3 + \dots + \alpha_q L^q$ ,  $\beta(L) = \beta_1 L^1 + \beta_2 L^2 + \beta_3 L^3 + \dots + \beta_p L^p$

Dicha función puede expresarse del siguiente modo :

$$h_t = \alpha_0 + \alpha(L)\varepsilon_t^2 + \beta(L)h_t \quad (3.11)$$

con  $\alpha_0 > 0, \alpha_i \geq 0, \gamma_i \geq 0, i = 1, 2, \dots, q, j = 1, 2, \dots, p$

Es importante observar que si consideramos :

1.  $p = 0$ , tenemos un proceso ARCH(q),  $h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2$
2.  $p = q = 0$ , tenemos un proceso ruido blanco.

$$h_t = \alpha_0 + \varepsilon_t$$

,

$$\varepsilon_t|\Psi_{t-1} \sim N(0, h_t)$$

Proceso GARCH(p,q) estacionario a partir de la ecuación 3.10 reemplazamos en la esperanza tenemos:

$$Var(\varepsilon_t) = E(\varepsilon_t^2) = E(E(\varepsilon_t^2|\Psi_{t-1})) = E(h_t) = \sigma_t^2$$

$$Var(\varepsilon_t) = \alpha_0 + \alpha_1 E(\varepsilon_{t-1}^2) + \alpha_2 E(\varepsilon_{t-2}^2) + \dots + \alpha_q E(\varepsilon_{t-q}^2) + \gamma_1 E(h_{t-1}) + \dots + \gamma_p E(h_{t-p})$$

$$Var(\varepsilon_t) = \alpha_0 + \alpha_1 Var(\varepsilon_{t-1}) + \alpha_2 Var(\varepsilon_{t-2}) + \dots + \alpha_q Var(\varepsilon_{t-q}) + \gamma_1 h_{t-1} + \gamma_2 h_{t-2} + \gamma_3 h_{t-3} + \dots + \gamma_p h_{t-p}$$

Si consideramos al proceso  $\varepsilon_t$ , estacionario, considerando la homocedasticidad en la serie  $Var(\varepsilon_t) = Var(\varepsilon_{t-k}), \forall k$   
Entonces

$$Var(\varepsilon_t) = \alpha_0 + \alpha_1 Var(\varepsilon_t) + \alpha_2 Var(\varepsilon_t) + \dots + \alpha_q Var(\varepsilon_t) + \gamma_1 Var(\varepsilon_t) + \gamma_2 Var(\varepsilon_t) + \dots + \gamma_p Var(\varepsilon_t)$$

$$Var(\varepsilon_t) - \alpha_1 Var(\varepsilon_t) - \alpha_2 Var(\varepsilon_t) - \dots - \alpha_q Var(\varepsilon_t) - \gamma_1 Var(\varepsilon_t) - \gamma_2 Var(\varepsilon_t) - \dots - \gamma_p Var(\varepsilon_t) = \alpha_0$$

$$Var(\varepsilon_t)(1 - \alpha_1 - \alpha_2 - \dots - \alpha_q - \gamma_1 - \dots - \gamma_p) = \alpha_0$$

$$Var(\varepsilon_t) = \frac{\alpha_0}{1 - \alpha_1 - \alpha_2 - \dots - \alpha_q - \gamma_1 - \dots - \gamma_p} = \sigma_t^2$$

$$Var(\varepsilon_t) = \frac{\alpha_0}{1 - [\alpha_1 + \alpha_2 + \dots + \alpha_q + \gamma_1 + \dots + \gamma_p]} = \sigma_t^2$$

Por lo tanto consideramos la siguiente restricción:

$$\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \gamma_j < 1 \quad (3.12)$$

Es importante mencionar que una señal de memoria larga en la serie se da cuando  $\gamma_j$  se aproxima a 1, claro esta se debe realizar las pruebas tanto estadísticas paramétricas semiparamétricas y no paramétricas para demostrar existencia de suficiente evidencia de persistencia en la serie.

Si incluimos la innovación unitaria  $v_t$  como la innovación de la varianza condicional las cuales son variables aleatorias independientes e idénticamente distribuidas estacionarias e incorrelacionadas serialmente (shock) :

$$v_t = \varepsilon_t^2 - h_t \quad (3.13)$$

Consideramos la Ecuación 3.8, entonces,  $\varepsilon_t^2 = h_t z_t^2$ , desarrollando:

$$\begin{aligned} v_t &= h_t z_t^2 - h_t \\ v_t &= h_t(z_t^2 - 1), \text{ recordando que } E(z_i, z_j) = 0, \forall i \neq j \text{ y también } z_t | \Psi_{t-1} \sim N(0, 1). \end{aligned}$$

Entonces

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \gamma_j h_{t-j} \quad (3.14)$$

Utilizando la ecuación 3.13 tenemos:  $v_t = \varepsilon_t^2 - h_t$   
 $h_t = \varepsilon_t^2 - v_t$

$$\varepsilon_t^2 - v_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \gamma_j (\varepsilon_t^2 - v_t)$$

$$\varepsilon_t^2 - v_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \gamma_j \varepsilon_{t-j}^2 - \sum_{j=1}^p \gamma_j v_{t-j}$$

$$\varepsilon_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \gamma_j \varepsilon_{t-j}^2 - \sum_{j=1}^p \gamma_j v_{t-j} + v_t$$

En el año 1986 Bollerslev y Engle, consideraron una clase de modelos GARCH(p,q) integrado denominado IGARCH(p,q), en donde la dependencia a largo plazo de los incrementos del proceso es permanente a diferencia de los modelos ARCH(q) y GARCH(p,q) en las cuales dichas persistencias desaparecen exponencialmente hacia cero.

En dicho modelo IGARCH(p,q) se cumple:

$$\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \gamma_j = 1$$

A diferencia del modelo GARCH(p,q)

$$\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \gamma_j < 1$$

En los modelos IGARCH(p,q), el impacto del cuadrado de los shocks :

$$v_{t-i} = \varepsilon_{t-i}^2 - h_{t-i} \quad (3.15)$$

para  $i > 0$ , sobre el cuadrado de los residuos es persistente. Por lo tanto en un modelo IGARCH(1,1), el modelo puede escribirse de la siguiente manera:

$$\begin{aligned} \varepsilon_t &= h_t^{1/2} z_t \\ h_t &= \alpha_0 + \beta_1 h_{t-1} + (1 - \beta_1) \varepsilon_{t-1}^2 \end{aligned} \quad (3.16)$$

Siendo  $\beta \in \langle 0, 1 \rangle$

Dicho modelo IGARCH(1,1) es estacionario y ergódico.

Entonces el modelo IGARCH(p,q) implica una persistencia infinita de la varianza condicional para los shocks en el cuadrado de los retornos, una persistencia infinita de los shocks de volatilidad.

Para el proceso FIGARCH(p,d,q), a partir de un proceso GARCH(p,q), puede ser expresado como un proceso ARMA(m,p) en el cuadrado de los residuos, entonces:

$$\varepsilon_{t-i}^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \gamma_j \varepsilon_{t-j}^2 - \sum_{j=1}^p \gamma_j v_{t-j} + v_t$$

$$\varepsilon_t^2 = \alpha_0 + \alpha(L)\varepsilon_t^2 + \beta(L)\varepsilon_t^2 - \beta(L)v_t + v_t$$

$$\varepsilon_t^2 - \alpha(L)\varepsilon_t^2 - \beta(L)\varepsilon_t^2 = \alpha_0 - \beta(L)v_t + v_t$$

$$(1 - \alpha(L) - \beta(L))\varepsilon_t^2 = \alpha_0 + (1 - \beta(L))v_t$$

$$v_t = \varepsilon_t^2 - h_t$$

Donde  $m = \max\{p, q\}$

Por lo tanto un proceso IGARCH(p,q) se expresa de la siguiente forma:

$$(1 - L)\Phi(L)\varepsilon_t^2 = \alpha_0 + [1 - \beta(L)]v_t$$

Donde:

$\Phi(L) = \sum_{i=1}^{m-1} \Phi_i L^i$ , La cual es de orden m-1

En este Modelo IGARCH(p,q) dicha persistencia permanente se refiere a la función de autocorrelación de la serie de tiempo, por otro lado en los modelos FIGARCH las caídas de la función de autocorrelación son hiperbólicas, por lo tanto tienden a cero las autocorrelaciones pero lentamente, la persistencia decae lentamente a un ritmo o grado la cual es interpretado por la estimación del parámetro fraccional "d",  $d \in (0, 1)$ .

El modelo FIGARCH(p,d,m) se define:

$$(1 - L)^d \Phi(L)\varepsilon_t^2 = \alpha_0 + (1 - \beta(L))v_t \quad (3.17)$$

Donde:

$$\Phi(L) = \sum_{i=1}^{m-1} \Phi_i L^i \quad (3.18)$$

la cual es de orden  $m-1$

El modelo FIGARCH(p,d,m), los shocks son exactamente la innovación en la representación ARMA del proceso FIGARCH(p,d,m), Así este shock

$$\varepsilon_t^2 = \alpha_0((1-L)^d\Phi(L))^{-1} + (1-\beta(L))((1-L)^d\Phi(L))^{-1}v_t \quad (3.19)$$

$v_t$  seria el shock y  $(1-\beta(L))((1-L)^d\Phi(L))^{-1}$  se interpretaría como la persistencia.

Este shock se interpretaría como el error de pronóstico de la varianza o también la variabilidad de la varianza no explicada

Si realizamos una analogía de la estructura FIGARCH con la estructura ARFIMA (p,d,q) un modelo lineal homocedástico al igual que en FIGARCH (p,d,q) con alta dependencia entre los datos en donde a diferencia que en el primero la segunda es el operador de la memoria larga sobre la media incondicional  $\mu$  de  $y_t$  del proceso, en el modelo FIGARCH la dependencia es no lineal.

El modelo ARFIMA introducido por Granger y Joyeux (1980), Granger(1981) y Hosking (1981) es definido :

$$(1-L)^d a(L)y_t = b(L)z_t \quad (3.20)$$

En donde  $a(L)$  y  $b(L)$ , son operadores de rezago de ordenes  $p$  y  $q$ , y  $Z$  es un proceso con media zero y serialmente incorrelacionada, la antipersistencia recae sobre la media, considerando que el parámetro de integracional fraccional "d" se encuentra entre el intervalo cerrado  $[-0.5, 0.5]$ , Análogamente al modelo FIGARCH (p,d,q), tendríamos al igual que en la ecuación 3.17  $(1-L)^d\Phi(L)\varepsilon_t^2 = \alpha_0 + (1-\beta(L))v_t$

Si a partir de la ecuación 3.13,  $v_t = \varepsilon_t^2 - h_t$ , realizamos un proceso inverso en 3.19, entonces tendríamos:

$$(1-L)^d\Phi(L)\varepsilon_t^2 = \alpha_0 + (1-\beta(L))(\varepsilon_t^2 - h_t)$$

$$(1-L)^d\Phi(L)\varepsilon_t^2 = \alpha_0 + (1-\beta(L))\varepsilon_t^2 - [1-\beta(L)]h_t$$

$$[1-\beta(L)]h_t = \alpha_0 + (1-\beta(L))\varepsilon_t^2 - (1-L)^d\Phi(L)\varepsilon_t^2$$

$$[1-\beta(L)]h_t = \alpha_0 + [1-\beta(L) - (1-L)^d\Phi(L)]\varepsilon_t^2$$

$$h_t = \alpha_0(1-\beta(L))^{-1} + [1 - (1-\beta(L))^{-1}\Phi(L)(1-L)^d]\varepsilon_t^2$$

$$h_t = \alpha_0[1-\beta(L)]^{-1} + \sum_i^\infty \lambda_i(L)^i \varepsilon_t^2$$

$$h_t = \alpha_0[1-\beta(L)]^{-1} + \lambda(L)\varepsilon_t^2$$

En donde,  $\lambda(L) = \lambda_1 L + \lambda_2 L^2 + \lambda_3 L^3 + \dots$ , con  $\lambda_k \geq 0$

para  $k = 1, 2, 3, \dots$

y satisfaciendo la nonegatividad de la función de la varianza condicionada, las cuales se estarán tratando en la siguiente sección.

Por lo tanto se considera como un proceso no estacionario debido a que un proceso autosimilar es no estacionario pero con incrementos estacionarios e dependientes y esta dependencia es no lineal debido a ser un movimiento browniano fraccionario con  $H(2H-1)n^{2H-2} > 0$  para  $H \in (0,5,1)$  existiendo una dependencia a largo plazo o persistencia del shock en la volatilidad (persistencia en la autocorrelación de las innovaciones) dichos incrementos tienden la misma dirección las trayectorias tienden a ir en la misma dirección.

$$d \in \begin{cases} (-0,5, 0) & \text{Decimos } \textit{estacionario antipersistente} \\ (0, 0,5) & \text{Decimos } \textit{estacionario de memoria larga} \\ (0,5, 1) & \text{Decimos } \textit{no estacionario de memoria larga con reversión a la media} \end{cases}$$

Es importante señalar que si  $d \in (-0,5, 0,5)$  se garantiza la estacionariedad del proceso. Por lo cual el proceso estocástico con incrementos estacionario y dependientes cuya dependencia es no lineal con índice de similaridad  $H \in (0,5, 1)$  es un proceso autosimilar con parámetro de memoria larga  $d \in (0, 0,5)$  con dichos incrementos los cuales tienden a la misma dirección y reversible a la media, pero si  $d \in (0,5, 1)$  entonces se define como un proceso con incrementos no estacionarios y incrementos dependientes con una menor velocidad en la disminución de la covarianza de los incrementos de los intervalos ajenos, una extensión del movimiento browniano fraccional (memoria larga como la persistencia de la autocorrelación de las innovaciones cuadráticas). define

$$\varepsilon_t | \Psi_{t-1} \sim N(0, h_t)$$

$$(1 - \alpha(L) - \beta(L))(1 - L)^d \varepsilon_t^2 = \alpha_0 + (1 - \beta(L))v_t$$

$$h_t = \alpha_0(1 - \beta(L))^{-1} + [1 - (1 - \beta(L))^{-1}\Phi(L)(1 - L)^d]\varepsilon_t^2$$

$$\text{con } d \in [0, 1]$$

### 3.1.4. Ruido Blanco Fraccionario

El modelo FIGARCH o GARCH con integración Fraccional es un modelo obtenido al reemplazar el operador de diferencias  $(1 - L)$  por el operador de diferencias Fraccional  $(1 - L)^d$ , donde  $d$  es una fracción entre  $0 < d < 1$ .

Así el modelo FIGARCH obtenido en consideración es:

$$(1 - \alpha(L) - \beta(L))(1 - L)^d \varepsilon_t^2 = \alpha_0 + (1 - \beta(L))v_t$$

De esta manera se obtiene un modelo que captura de mejor manera la dependencia temporal observada de la volatilidad de los mercados financieros.

Entonces el termino  $(1 - L)^d$  es definido como ruido Blanco fraccionario, que de acuerdo con Lombard y Gallo (2005), si se trata a  $L$  como un escalar puede expresarse el binomio de Newton al realizar la expansión mediante la serie de Mclaurin se obtiene:

$$(1 - L)^d = \sum_{k=0}^d \binom{d}{k} (-1)^k (L)^k = \binom{d}{0} (-L)^0 + \binom{d}{1} (-L)^1 + \binom{d}{2} (-L)^2 + \binom{d}{3} (-L)^3 + \binom{d}{4} (-L)^4 + \dots$$

$$(1 - L)^d = \frac{d!}{d!} L^0 - \frac{d!}{(d-1)!1!} L + \frac{d!}{(d-2)!2!} L^2 - \frac{d!}{(d-3)!3!} L^3 + \frac{d!}{(d-4)!4!} L^4 - \dots$$

$$(1-L)^d = 1 - \frac{d}{1!} L + \frac{(d-1)d}{2!} L^2 - \frac{(d-2)(d-1)d}{3!} L^3 + \frac{(d-3)(d-2)(d-1)d}{4!} L^4 - \dots$$

Donde  $d$  puede tomar un valor no entero. Dado que la función factorial está definida solo para números naturales, tenemos que redefinir los conceptos binomiales utilizando la función gamma.

Entonces:

$$\binom{d}{k} = \frac{d!}{k!(d-k)!} = \frac{\Gamma(d+1)}{\Gamma(k+1)\Gamma(d-k+1)}$$

Donde se define a la función  $\Gamma$  de Legendre:

$$\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt$$

La cual es una extensión del factorial a los números complejos y cumple para los enteros positivos:  $\Gamma(i) = (i-1)!$

Recurriendo nuevamente a la expansión como serie de Mclaurin se tiene:

$$(1 - L)^d = \sum_{k=0}^{\infty} \frac{\Gamma(d+1)}{\Gamma(k+1)\Gamma(d-k+1)} (-L)^k$$

Por lo tanto el parámetro  $d \in \langle 0, 1 \rangle$ , entonces si  $d$  es un número natural y volvemos a expandir mediante la serie de McLarin, se obtienen los coeficientes de  $L^K$ .

$$\Psi_k = \frac{(k + d - 1)!}{k!(d - 1)!}$$

O en recursiva,  $\Psi_k = \Psi_{k-1} \frac{(k+d-1)}{k}$ , con  $\Psi_0 = 1$ .  
Entonces se puede escribir de la siguiente manera:

$$\Psi_k = \frac{\Gamma(k + d)}{\Gamma(k + 1)\Gamma(d)}$$

Por tanto se puede expresar de la siguiente forma:

$$(1 - L)^d = \sum_{k=0}^{\infty} \frac{\Gamma(k - d)}{\Gamma(k + 1)\Gamma(-d)} (L)^k \quad (3.21)$$

### 3.1.5. La nonegatividad de la varianza condicionada

Si definimos la ecuación de la varianza condicionada ya desarrollada en apartados anteriores:

$$h_t = \alpha_0(1 - \beta(L))^{-1} + [1 - (1 - \beta(L))^{-1}\Phi(L)(1 - L)^d]\varepsilon_t^2 \quad (3.22)$$

El proceso FIGARCH proporciona una gran flexibilidad en el modelado de la dependencia temporal de la varianza condicional y toma en consideración como casos especiales a GARCH( $d=0$ ) y el IGARCH( $d=1$ ). En relación a la estacionariedad se considera débilmente estacionaria cuando  $d=0$ , es decir es un proceso GARCH de memoria corta o antipersistencia, Pero es importante considerar que según Baillie, Bollerslev y Mikkelsen (1996) consideran que el modelo es estrictamente estacionario y ergódico al igual que IGARCH. Pero existe en el modelo FIGARCH una cierta limitación en cuanto a la no negatividad de la varianza, es decir los parámetros deben cumplir restricciones que no necesariamente son fáciles de satisfacer, los cuales se tratan en mayor profundidad a continuación.

Si consideramos un modelo FIGARCH (1, $d$ ,1), una condición sugerida para satisfacer la nonegatividad de la varianza condicionada fue dada por Baillie (1996), Bollerslev y Mikkelsen (1996) utilizando los coeficientes de FIGARCH (1, $d$ ,1), dicho modelo definida en la ecuación siguiente.

$$h_t = \alpha_0(1 - \beta(1))^{-1} + [1 - (1 - \beta_1(L))^{-1}(1 - \Phi_1(L))(1 - L)^d]\varepsilon_t^2$$

Donde:

$$\lambda(L) = \lambda_1 L + \lambda_2 L^2 + \lambda_3 L^3 + \dots = 1 - [(1 - \beta_1(L))^{-1}(1 - \Phi_1 L)(1 - L)^d] \quad (3.23)$$

Por lo tanto:

$$\lambda_k = \beta_1 \lambda_{k-1} + \left[ \frac{k - 1 - d}{k} - \Phi_1 \right] \delta_{d,k-1}$$



$$k = 2, 3, \dots$$

$$\delta_{d,k} = \delta_{d,k-1} \frac{k-1-d}{k}, \text{ entonces}$$

$$\delta_d(L) = \sum_{k=1}^{\infty} \delta_{d,k} L^k, \text{ con } \delta_{d,0} = 1$$

A partir de lo mencionado podemos llegar a las desigualdades sugeridas por Baillie(1996) y Chung(2001), entonces:

$$\beta_1 - d \leq \Phi_1 \leq \frac{2-d}{3}$$

$$d(\Phi_1 - \frac{1-d}{2}) \leq \beta_1(d - \beta_1 + \Phi_1)$$

$$0 \leq \Phi \leq \beta \leq d \leq 1$$

Dichas condiciones dadas por Baille(1996) y Chung(2001), no son equivalentes y puede que un conjunto de valores de los parámetros satisfagan una desigualdad y no la otra.

En el año 2006 Conrad y Haag desarrollaron un conjunto de condiciones para garantizar la nonegatividad de la varianza condicional en todas las situaciones.

**Teorema 3.1.1** (Conrad y Haag, 2006). Sea  $\varrho_j$  los coeficientes y  $\mathfrak{f}_j$  la función del parámetro de integración fraccional  $d$ , Asi tengo,  $\varrho_j = \mathfrak{f}_j \varrho_{j-1} = \prod_{i=1}^j \mathfrak{f}_i$  con  $\mathfrak{f}_j = \frac{j-1-d}{j}$  para  $j=1,2,\dots$  y  $\varrho_0 = 1$  entonces las condiciones son:  $\psi_1 = d + \phi_1 - \beta_1$  y  $\beta_1 \psi_{i-1} + (\mathfrak{f}_i - \phi_1)(-\varrho_{i-1}), i \geq 2$  alternativamente  $\psi_i = \beta_1^2 \psi_{i-2} + [\beta_1(\mathfrak{f}_{i-1} \phi_1) + (\mathfrak{f}_i - \phi_1) \mathfrak{f}_{i-1}](-\varrho_{i-2})$  para todo  $i \geq 3$ .

### 3.1.6. La función Impulso respuesta

La función impulso respuesta se utiliza para analizar y cuantificar el impacto de un shock sobre la volatilidad de una serie de memoria larga, es decir una serie caótica muy sensible a mínimos cambios en las condiciones iniciales, teniendo en cuenta que dicho impacto tiende a cero pero en este caso de manera hiperbólica, entonces dicha persistencia de la varianza condicionada es caracterizada en términos de los coeficientes de dicha función:

$$\gamma_k = \frac{\partial E(\varepsilon_{t+k}^2 | \Psi_{t-1})}{\partial \nu_t} - \frac{\partial E(\varepsilon_{t+k-1}^2 | \Psi_{t-1})}{\partial \nu_t}$$

$$\gamma_k = \frac{\partial h_{t+k}}{\partial \nu_t} - \frac{\partial h_{t+k-1}}{\partial \nu_t}$$

Si a partir de  $v_t = \varepsilon_t^2 - h_t$ , Entonces tendríamos:

$$\gamma_k = \frac{\partial h_{t+k}}{\partial \nu_t} - \frac{\partial h_{t+k-1}}{\partial \nu_t} = L(\lambda) \left\{ \frac{\partial E(\varepsilon_{t+k}^2 | \Psi_{t-1})}{\partial \nu_t} - \frac{\partial E(\varepsilon_{t+k-1}^2 | \Psi_{t-1})}{\partial \nu_t} \right\}$$

Entonces para encontrar los coeficientes de la función impulso respuesta se tendrá lo siguiente

$$h_t = \alpha_0(1 - \beta(L))^{-1} + [1 - (1 - \beta(L))^{-1}\Phi(L)(1 - L)^d]\varepsilon_t^2$$

$$\varepsilon_t^2 - v_t = \alpha_0(1 - \beta(L))^{-1} + [1 - (1 - \beta(L))^{-1}\Phi(L)(1 - L)^d]\varepsilon_t^2$$

$$\varepsilon_t^2 - [1 - (1 - \beta(L))^{-1}\Phi(L)(1 - L)^d]\varepsilon_t^2 = \alpha_0(1 - \beta(L))^{-1} + v_t$$

$$\varepsilon_t^2(1 - [1 - (1 - \beta(L))^{-1}\Phi(L)(1 - L)^d]) = \alpha_0(1 - \beta(L))^{-1} + v_t$$

$$\varepsilon_t^2(-(1 - \beta(L))^{-1}\Phi(L)(1 - L)^d) = \alpha_0(1 - \beta(L))^{-1} + v_t$$

$$\varepsilon_t^2 = -[\alpha_0(1 - \beta(L))^{-1} + v_t][(1 - \beta(L))\Phi(L)^{-1}(1 - L)^{-d}]$$

$$\varepsilon_t^2 = -[\alpha_0(1 - \beta(L))^{-1}][(1 - \beta(L))\Phi(L)^{-1}(1 - L)^{-d}] - v_t[(1 - \beta(L))\Phi(L)^{-1}(1 - L)^{-d}]$$

$$\varepsilon_t^2 = -[\alpha_0(\Phi(L))^{-1}][\Phi(L)(\Phi(L))^{-1}(1 - L)^{-d}] - v_t[\beta_k(\Phi(L))^{-1}(1 - L)^{-d}]$$

$$\varepsilon_t^2 = -[\alpha_0(\Phi(L))^{-1}][(1-L)^{-d}] - v_t[\beta_k(\Phi(L))^{-1}(1-L)^{-d}]$$

$$\frac{\partial E(\varepsilon_{t+k}^2 | \Psi_{t-1})}{\partial \nu_t} = -[\beta_k(\Phi(L))^{-1}(1-L)^{-d}]L(\lambda)$$

$$\frac{\partial E(\varepsilon_{t+k}^2 | \Psi_{t-1})}{\partial \nu_t} = -[\beta_k(\Phi(L))^{-1}(1-L)^{-d}][1 - (1 - \beta(L))^{-1}\Phi(L)(1-L)^d]$$

$$\frac{\partial E(\varepsilon_{t+k}^2 | \Psi_{t-1})}{\partial \nu_t} = -[\beta_k[1 - (1 - \beta(L))^{-1}\Phi(L)(1-L)^d](\Phi(L))^{-1}(1-L)^{-d}]$$

$$= -\beta_k(\Phi(L))^{-1}(1-L)^{-d} + \beta_k(1-L)^{-d}$$

$$\frac{\partial E(\varepsilon_{t+k-1}^2 | \Psi_{t-1})}{\partial \nu_t} = -\beta_{k-1}(\Phi(L))^{-1}(1-L)^{-d} + \beta_k(1-L)^{-d}$$

Por lo tanto.

$$\gamma_k = -\beta_k(\Phi(L))^{-1}(1-L)^{-d} + \beta_{k-1}(\Phi(L))^{-1}(1-L)^{-d} - \beta_k(1-L)^{-d}$$

$$\gamma_k = (\beta_{k-1} - \beta_k)(\Phi(L))^{-1}(1-L)^{-d} \quad (3.24)$$

Lo que demuestra que en el modelo Figarch(1,d,1), los coeficientes de la función impulso respuesta no dependen del tiempo  $t$ , los cuales indican la persistencia de la varianza condicionada (Baillie, 1996).

El término ruido blanco fraccionario de la ecuación 3.21 también puede expresarse en términos de una función hipergeométrica de Gauss de la siguiente forma:

$$(1-L)^d = F(-d, 1, 1; L) \quad (3.25)$$

$$F(-d, 1, 1; L) = \sum_{k=0}^{\infty} \frac{(-d)_k (1)_k}{k! (1)_k} L^k = \sum_{k=0}^{\infty} \frac{(-d)_k}{k!} L^k$$

Si  $(a)_k$  se denomina símbolo de Pochhammer.

Donde:  $(a)_k \begin{cases} 1 & \text{si; } k=0 \\ a(a+1)\dots(a+k-1) & \text{si ; } k=1,2,3,\dots \end{cases}$

$$\sum_{k=0}^{\infty} \frac{\Gamma(-d+k)}{\Gamma(-d)k!} L^k = \sum_{k=0}^{\infty} \frac{\Gamma(k-d)}{\Gamma(k+1)\Gamma(-d)} L^k$$

Se demuestra la igualdad de la ecuacion 3.21

Por lo tanto si combinamos las ecuaciones 3.24 con 3.25, es decir tomando en cuenta la persistencia del shock  $(1 - \beta(L))(\Phi(L))^{-1}(1 - L)^{-d}$

$$\gamma_k = (1 - \beta(L))(\Phi(L))^{-1}(1 - L)^{-d}$$

Es equivalente a:

$$\gamma_k = (1 - \beta(L))(\Phi(L))^{-1}F(-d, 1, 1; L) \quad (3.26)$$

Por lo tanto la función impulso respuesta acumulativa se expresa:

$$\gamma_1 = \varphi(0) + \varphi(1) + \varphi(2) + \dots$$

$$\gamma_1 = \lim_{s \rightarrow \infty} \sum_{i=0}^s \varphi(i)$$

Si definimos la ecuación 3.25 como  $F(d-1, 1, 1, L) = (1 - L)^{1-d}$ , para  $0 \leq d < 1$ , entonces se presentan los siguientes casos:

Caso  $d = 0$  se genera un GARCH(p,q) memoria corta relacionado con el proceso gaussiano con incrementos estacionario e independientes con índice de similaridad  $H = \frac{1}{2}$ , (movimiento browniano)

La función impulso respuesta acumulativa es:

$$\gamma_1 = (1 - \beta_1(1))(\Phi_1(1))^{-1}F(d-1, 1, 1; 1)$$

$$\gamma_1 = (1 - \beta_1(1))(\Phi_1(1))^{-1}(1 - L)^{1-d}$$

$$\gamma_1 = (1 - \beta_1(1))(\Phi_1(1))^{-1}(1 - 1)^{1-0}$$

$$\gamma_1 = 0$$

$$F(d-1, 1, 1; 1) = 0$$

$$\lim_{s \rightarrow \infty} \sum_{i=0}^s \varphi(i) = 0 \quad (3.27)$$

La varianza condicionada se estabiliza automáticamente la covarianza de intervalos ajenos desaparece exponencialmente.

Caso  $d = 1$  se genera un IGARCH(p,q) memoria persistente y d raíces unitarias

La función impulso respuesta acumulativa es:

$$\gamma_1 = (1 - \beta_1(1))(\Phi_1(1))^{-1}F(d - 1, 1, 1; 1)$$

$$\gamma_1 = (1 - \beta_1(1))(\Phi_1(1))^{-1}(1 - L)^{1-d}$$

$$\gamma_1 = (1 - \beta_1(1))(\Phi_1(1))^{-1}(1 - L)^{1-1}$$

$$\gamma_1 = (1 - \beta_1(1))(\Phi_1(1))^{-1}$$

$$F(d - 1, 1, 1; 1) = 1$$

$$\lim_{s \rightarrow \infty} \sum_{i=0}^s \varphi(i) = (1 - \beta_1(1))(\Phi_1(1))^{-1} \quad (3.28)$$

La persistencia es indefinida, el efecto de un shock persiste indefinidamente.

Caso  $0 < d < 1$  se genera un *FIGARCH*(m, d, q) memoria larga relacionado con el proceso con incrementos estacionarios y dependientes o no estacionarios y dependientes tales situaciones determinados por d con índice de similaridad  $H > \frac{1}{2}$ .

La función impulso respuesta acumulativa es:

$$\gamma_1 = 0$$

$$F(d - 1, 1, 1; 1) = 0$$

$$\lim_{s \rightarrow \infty} \sum_{i=0}^s \varphi(i) = 0 \quad (3.29)$$

La persistencia desaparece hiperbólicamente tiende a cero al igual que garch pero la diferencia es la velocidad de las disminuciones de las autocorrelaciones y la serie revierte a su media.

Caso  $d > 1$

La función impulso respuesta acumulativa es:

$$\gamma_1 = (1 - \beta_1(1))(\Phi_1(1))^{-1}F(d - 1, 1, 1; 1)$$

$$\gamma_1 = (1 - \beta_1(1))(\Phi_1(1))^{-1}(1 - L)^{1-d}$$

$$\gamma_1 = \frac{(1 - \beta_1(1))(\Phi_1(1))^{-1}}{(1 - 1)^{1-d}}$$

$$\gamma_1 = \infty$$

$$F(d - 1, 1, 1; 1) = \infty$$

$$\lim_{s \rightarrow \infty} \sum_{i=0}^s \varphi(i) = \infty \quad (3.30)$$

La varianza condicionada es explosiva e irreal.

### 3.1.7. Estimación y Pronóstico

Generalmente para la estimación de los parámetros del modelo FIGARCH, se utiliza la estimación por máxima verosimilitud, siendo dicho método la que genera estimaciones consistentes y eficientes para muestras asintóticas considerando la normalidad de la innovación unitaria  $z_t \sim N(0, 1)$ .

Sabemos que el espacio de parámetros a estimar es:

$$\theta' = [\alpha_0, d, \beta_1, \beta_2, \beta_3, \dots, \beta_p, \Phi_1, \Phi_2, \Phi_3, \dots, \Phi_q]$$

Considerando una estructura FIGARCH (p,d,q), a partir de una muestra de variables aleatorias independientes e idénticamente distribuidas conocidas,  $[\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots, \varepsilon_T]$

$$\hat{\theta} \sim N(\theta, (I_{\theta\theta})^{-1}), \quad (3.31)$$

$I_{\theta\theta}$ , es la matriz de información de Fisher

La estimación máxima verosimilitud de  $\theta_1, \theta_2, \theta_3, \dots, \theta_p$  son los valores  $\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \dots, \hat{\theta}_p$ , que maximizan la verosimilitud, entonces la estimación de máxima verosimilitud de es denotado por

$$\theta_i = T_i(\varepsilon), i = 1, 2, 3, \dots, p, \quad (3.32)$$

Y la estimación de máxima verosimilitud de  $\vartheta = g(\theta_1, \theta_2, \theta_3, \dots, \theta_p)$  es  $\hat{\vartheta} = g(\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \dots, \hat{\theta}_p)$ , Bajo las condiciones generales la estimación de máxima verosimilitud de  $\vartheta$  es asintóticamente normal, asintóticamente insesgado y eficiente. Por lo tanto asintóticamente para n,  $\hat{\vartheta} \sim N(\vartheta, CRLB(\vartheta))$  se cumple, aproximadamente .Esto es llamada una distribución asintótica de  $\vartheta$ .

Si consideramos,  $\varepsilon_t | \Psi_{t-1} \sim N(0, h_t)$ , entonces la función de logaritmo de verosimilitud para la t-ésima observacion es:

$$f(\varepsilon_t | \Psi_{t-1}) = \frac{1}{\sqrt{2\pi h_t}} e^{-\frac{1}{2} \frac{\varepsilon_t^2}{h_t}}$$

$$f(\varepsilon_t | \Psi_{t-1}) = (2\pi h_t)^{-\frac{1}{2}} e^{-\frac{1}{2} \frac{\varepsilon_t^2}{h_t}}$$

$$\ln f(\varepsilon_t | \Psi_{t-1}) = -\frac{1}{2} \ln(2\pi h_t) - \frac{1}{2} \frac{\varepsilon_t^2}{h_t}$$

$$\ln f(\varepsilon_t | \Psi_{t-1}) = -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln h_t - \frac{1}{2} \frac{\varepsilon_t^2}{h_t}$$

Para las T observaciones

$$f(\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots, \varepsilon_T, \theta) = \prod_{t=1}^T \frac{1}{\sqrt{2\pi h_t}} e^{-\frac{1}{2} \frac{\varepsilon_t^2}{h_t}}$$

$$\ln f(\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots, \varepsilon_T, \theta) = -\frac{T}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^T (\ln h_t) - \frac{1}{2} \sum_{t=1}^T \frac{\varepsilon_t^2}{h_t}$$

Sabemos que:  $\varepsilon_t = h_t^{1/2} z_t$ , Entonces:

$$\varepsilon_t^2 = h_t z_t^2$$

$$\ln f(\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots, \varepsilon_T, \theta) = -\frac{T}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^T \ln(h_t) - \frac{1}{2} \sum_{t=1}^T \frac{h_t z_t^2}{h_t}$$

$$\ln f(\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots, \varepsilon_T, \theta) = -\frac{T}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^T \ln(h_t) - \frac{1}{2} \sum_{t=1}^T z_t^2$$

$$\ln f(\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots, \varepsilon_T, \theta) = -\frac{T}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^T (\ln h_t + z_t^2)$$

$$\ln(\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots, \varepsilon_T, \theta) = -\frac{T}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^T (\ln h_t + \varepsilon_t^2 h_t^{-1})$$

Pero en la mayoría de las series de data financiera de alta frecuencia, la innovación unitaria  $z_t = \varepsilon_t^2 h_t^{-1}$ , tiene un exceso de kurtosis, leptocurtica y con colas más pesadas que una normal estándar, de tal manera el reducir el exceso de kurtosis en la distribución de los residuos de la varianza condicionada llevaría consigo una utilización de alguna distribución más apropiada, como la de Pareto- levy, t-student, skewed student distribution. En este contexto un procedimiento más

robusto, sería la de quasimáxima verosimilitud(QMLE) , dicho procedimiento fue discutido por Weiss(1986), Bollerslev y Woodridge(1992), generando mejores resultados para la, inferencia.

$$\sqrt{T}(\hat{\theta}_T - \theta_0) \longrightarrow N(0, A(\theta_0)^{-1} B(\theta_0) A(\theta_0)^{-1})$$

donde:  $A(\theta_0)$  es la hessiana y  $B(\theta_0)$  es la gradiente

Siendo esta la QMLE obtenida de maximizar la función de logaritmo de verosimilitud anterior, por lo cual  $\hat{\theta}_T$  es consistente y con normalidad asintótica para un modelo FIGARCH(1,d,0).



En la actualidad aún no se tiene una prueba que demuestre la consistencia y normalidad asintótica de la estimación por quasimáxima verosimilitud para el modelo general de FIGARCH (p,d,q). En dicho procedimiento de la estimación

de los parámetros del modelo, se estima primero el parámetro de integración fraccional  $d$ , para luego a partir ella se estiman los demás parámetros. Ahora si consideramos el problema del pronóstico del modelo FIGARCH, se realiza a partir de la ecuación de la varianza condicionada (Tayefi and Ramanathan, 2012):

$$h_t = \alpha_0(1 - \beta(L))^{-1} + [1 - (1 - \beta(L))^{-1}\Phi(L)(1 - L)^d]\varepsilon_t^2$$

$$h_{t+1} = \alpha_0(1 - \beta(1))^{-1} + \lambda(L)\varepsilon_{t+1}^2$$

$$h_t(1) = \alpha_0(1 - \beta(1))^{-1} + \lambda_1\varepsilon_t^2 + \lambda_2\varepsilon_{t-1}^2 + \dots$$

Entonces se dan el segundo paso para la prediccion:

$$h_t(2) = \alpha_0(1 - \beta(1))^{-1} + \lambda_1h_t(1) + \lambda_2\varepsilon_t^2 + \dots$$

En general los l-pasos para la prediccion:

$$h_t(l) = \alpha_0(1 - \beta(1))^{-1} + \lambda_1h_t(l-1) + \dots + \lambda_{l-1}h_t(1) + \lambda_l\varepsilon_t^2 + \lambda_{l+1}\varepsilon_{t-1}^2 + \dots$$

Entonces se tiene:

$$h_t(l) \approx \alpha_0(1 - \beta(1))^{-1} + \sum_{i=1}^{l-1} \lambda_i h_t(l-i) + \sum_{k=0}^M \lambda_{l+k} \varepsilon_{t-k}^2$$

## **3.2. Población y tamaño de muestra**

### **3.2.1. Población**

La población estuvo determinada por el conjunto de datos del tipo de cambio nominal del sol con respecto al dólar americano

### **3.2.2. Muestra**

El tamaño de muestra que nos sirvió de insumo para la estimación del modelo Figarch y los modelos heterocedásticos fue con un total de 20 años (1997- 2017) analizados diariamente, siendo un total de 5266 observaciones, los cuales fueron extraídos de la base de datos del Banco Central de Reserva del Perú.

# Capítulo 4

## Resultados

### 4.1. Técnicas de recolección de datos

Para la presente investigación se consideraron datos diarios de la serie estadística del Banco Central de Reserva del Perú (BCRP) de un periodo inicial del año 1997 al 2017, para el cual se considero desde el mes de enero de 1997 al mes de marzo del 2017. Con lo cual la estimación del modelo se dio a partir de una muestra de 5266 observaciones.

### 4.2. Técnicas para el procesamiento de la Información

#### 4.2.1. Análisis Estadístico

Con el objetivo de poder entender la dinámica de los retornos del tipo de cambio, se realizará un análisis estadístico y grafico de dicha serie. Teniendo en consideración que todos los cálculos del análisis descriptivo fueron hechas en el programa estadístico Oxmetrics y Gretel, es importante mencionar que debido a 229 observaciones missing se realizó un proceso de imputación de datos por el metodo de medianas de puntos adyacentes con numero de amplitud de puntos adyacentes 3 con el objetivo de no distorcionar la verdadera dinámica del comportamiento de la serie para este fin me ayude del software SPSS.

A continuación en la figura 4-1, se presenta la serie tipo de cambio, en la figura 4-2 sus retornos porcentuales y en las figuras 4-3 y 4-4, sus respectivas transformaciones (retorno al cuadrado y valor absoluto).

A primera vista se observa una no estacionariedad en media de la serie de tipo de cambio mientras que los retornos porcentuales parecen indicar estacionariedad por lo menos en media, los cuales también me indican la durabilidad de la volatilidad en la serie en el periodo de análisis por otro lado para analizar la dependencia temporal no lineal en los retornos porcentuales, se empieza a analizar la serie de retornos al cuadrado y en valor absoluto, indicando heterocedasticidad, es decir

Figura 4.1: Serie tipo de cambio

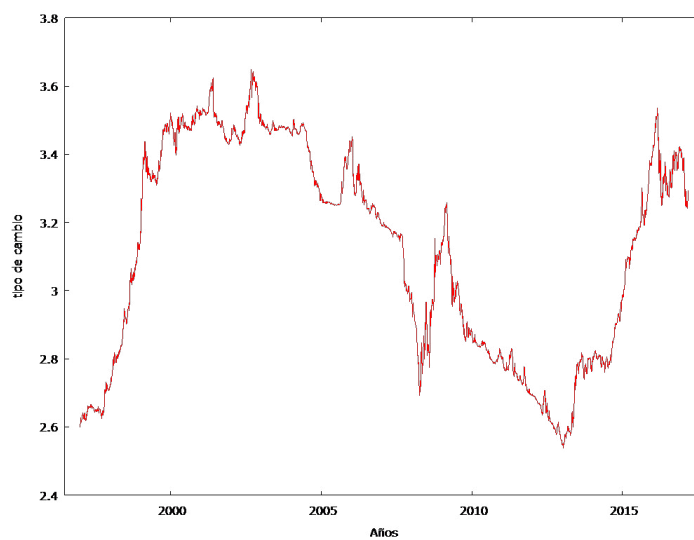


Figura 4.2: Retorno porcentual de la serie

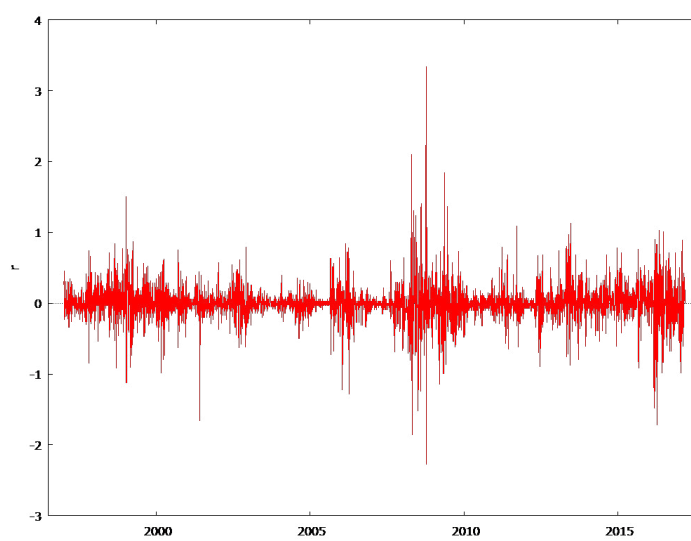


Figura 4.3: Cuadrado de los retornos porcentuales de la serie

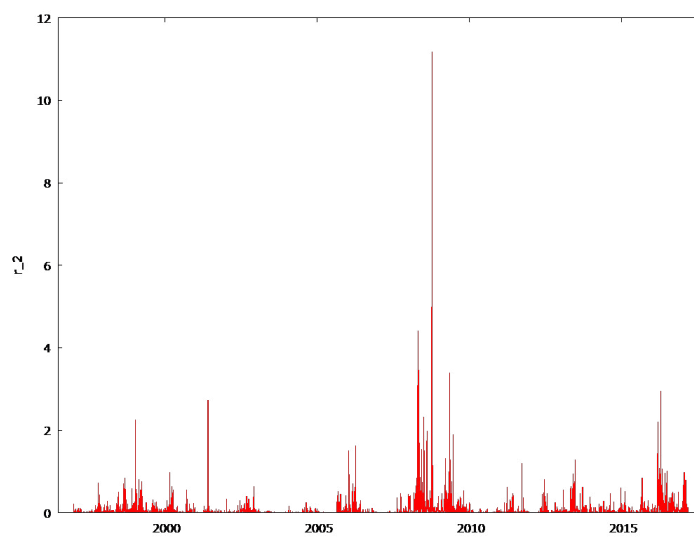
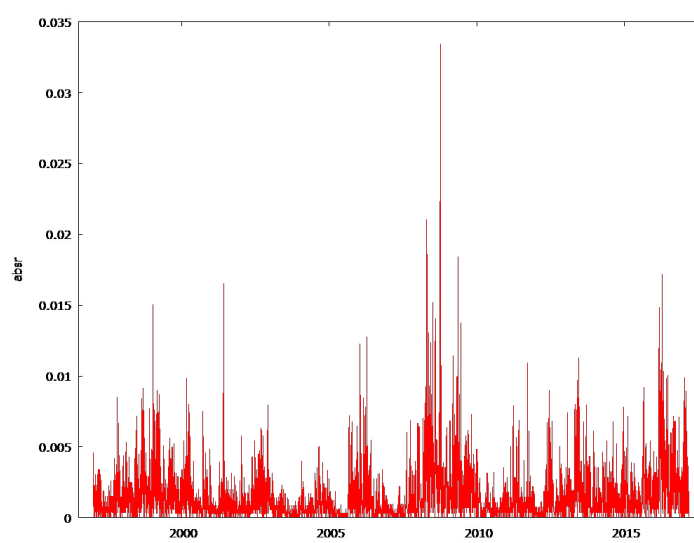


Figura 4.4: Valor absoluto de los retornos de la serie



la volatilidad es cambiante en el tiempo y me induce a pensar la existencia de la volatilidad por agrupamientos, grandes y pequeños cambios tienden a estar agrupados. Por otro lado la serie de retornos me indican una distribución leptocurtica una excesiva concentración alrededor de la media y se detectan outliers con respecto a la distribución normal.

Cuadro 4.1: Resumen descriptivo de los retornos porcentuales

| Principales Estadística |           |
|-------------------------|-----------|
| Media                   | 0.0048146 |
| Mediana                 | 0         |
| Mínimo                  | -2.27783  |
| Máximo                  | 3.34207   |
| Desv. Típica.           | 0.250518  |
| Asimetría               | 0.372755  |
| Exc. de curtosis        | 15.5737   |
| Porc. 5 %               | -0.354292 |
| Porc. 95 %              | 0.387921  |
| Rango IQ                | 0.179118  |

Fuente: Elaboración Propia.

La serie de retornos porcentuales me indican en primer lugar el retorno medio es de 0.0048146 por ciento para el periodo bajo análisis con una desviación de 0.250518 por ciento, el retorno máximo observado para este periodo es de 3.34207 por ciento, mientras que el mínimo es de -2.27783 por ciento , en las medidas de forma de la distribución es notable un exceso de kurtosis es decir la serie es leptocúrtica, por lo tanto tenemos una excesiva concentración de observaciones alrededor de la media con colas pesadas en comparación con la distribución gaussiana y se ve un número importante de outliers relativo a la distribución normal como se indica en el cuadro 4-1.

En el cuadro 4-2, es notable la no estacionariedad de la serie de los retornos porcentuales, rechazandose la hipótesis nula de estacionariedad en media, entonces se tiene una correlación serial, debido a que la probabilidad de la prueba de Ljung Box es menor al nivel de significancia del 5 por ciento.

A continuación en el cuadro 4-3, se evidencia la existencia de correlación serial en los retornos cuadráticos por lo tanto se detecta la posibilidad de memoria larga no solo en varianza sino también en media pero aún se necesitan más pruebas.

En la Figura 4-5, se muestra la prueba de jarque Bera(Apendice 3) la cual tiene una distribución chi-cuadrado con 2 grados de libertad, realiza la comparación entre una distribución simétrica y mesocúrtica con la serie de retornos porcentuales, al ser la probabilidad menor al 5 por ciento, entonces se rechaza la hipótesis nula de normalidad, las cuales son confirmados por la kurtosis de 18.57367 (leptocúrtica).

Cuadro 4.2: Autocorrelacion de los retornos porcentuales  
Función de autocorrelación para r (Utilizando desviación típica  $1/T^{0.5}$ )

| RETARDO | FAC     | FACP    | Estad-Q. | [valor p] |
|---------|---------|---------|----------|-----------|
| 1       | 0.1438  | 0.1438  | 108.8637 | [0.000]   |
| 2       | -0.0001 | -0.0212 | 108.8638 | [0.000]   |
| 3       | 0.037   | 0.041   | 116.0965 | [0.000]   |
| 4       | 0.0075  | -0.004  | 116.3898 | [0.000]   |
| 5       | 0.0118  | 0.0126  | 117.1258 | [0.000]   |
| 6       | 0.0358  | 0.0316  | 123.8784 | [0.000]   |
| 7       | -0.0051 | -0.015  | 124.0135 | [0.000]   |
| 8       | 0.0152  | 0.0188  | 125.2317 | [0.000]   |
| 9       | 0.0132  | 0.0056  | 126.1573 | [0.000]   |
| 10      | -0.0386 | -0.0412 | 134      | [0.000]   |
| 11      | 0.0101  | 0.0208  | 134.5334 | [0.000]   |
| 12      | 0.0255  | 0.0184  | 137.9635 | [0.000]   |
| 13      | 0.0204  | 0.0183  | 140.1693 | [0.000]   |
| 14      | 0.0079  | 0.0007  | 140.5018 | [0.000]   |
| 15      | -0.0069 | -0.0093 | 140.7501 | [0.000]   |
| 16      | -0.0184 | -0.0151 | 142.5351 | [0.000]   |
| 17      | -0.011  | -0.0096 | 143.1698 | [0.000]   |
| 18      | 0.0319  | 0.0356  | 148.562  | [0.000]   |
| 19      | 0.0404  | 0.0321  | 157.1955 | [0.000]   |
| 20      | 0.0764  | 0.0672  | 188.0929 | [0.000]   |
| 21      | 0.0534  | 0.0335  | 203.1698 | [0.000]   |
| 22      | 0.0071  | -0.0039 | 203.4396 | [0.000]   |
| 23      | 0.0109  | 0.0077  | 204.0652 | [0.000]   |
| 24      | 0.0043  | -0.0049 | 204.1647 | [0.000]   |
| 25      | 0.0153  | 0.0124  | 205.4003 | [0.000]   |
| 26      | 0.006   | -0.0054 | 205.5892 | [0.000]   |
| 27      | -0.0244 | -0.0284 | 208.7398 | [0.000]   |
| 28      | -0.0069 | 0.0016  | 208.995  | [0.000]   |
| 29      | -0.0098 | -0.0111 | 209.4999 | [0.000]   |
| 30      | -0.0109 | -0.0033 | 210.1328 | [0.000]   |
| 31      | -0.0061 | -0.0064 | 210.3303 | [0.000]   |
| 32      | 0.0019  | -0.0014 | 210.3494 | [0.000]   |
| 33      | -0.007  | -0.0091 | 210.6076 | [0.000]   |
| 34      | 0.0034  | 0.0045  | 210.6703 | [0.000]   |
| 35      | -0.0239 | -0.0211 | 213.6938 | [0.000]   |
| 36      | -0.0115 | -0.001  | 214.3904 | [0.000]   |
| 37      | -0.0018 | -0.0019 | 214.4067 | [0.000]   |

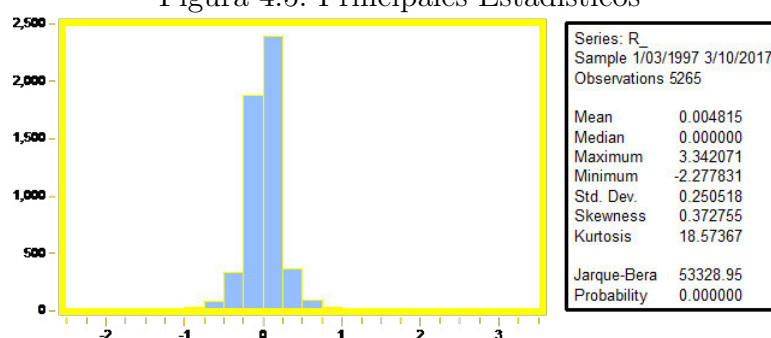
Cuadro 4.3: Autocorrelación de los retornos porcentuales al cuadrado

Función de autocorrelación para  $r^2$   
Utilizando desviación típica  $1/T^{0,5}$

| RETARDO | FAC    | FACP    | Estad-Q.  | [valor p] |
|---------|--------|---------|-----------|-----------|
| 1       | 0.3212 | 0.3212  | 543.3285  | [0.000]   |
| 2       | 0.3508 | 0.2762  | 1191.8782 | [0.000]   |
| 3       | 0.2879 | 0.142   | 1628.7012 | [0.000]   |
| 4       | 0.1445 | -0.0518 | 1738.7315 | [0.000]   |
| 5       | 0.1722 | 0.0407  | 1895.1334 | [0.000]   |
| 6       | 0.0955 | -0.0103 | 1943.175  | [0.000]   |
| 7       | 0.1599 | 0.1012  | 2078.0214 | [0.000]   |
| 8       | 0.0974 | -0.0008 | 2128.0395 | [0.000]   |
| 9       | 0.103  | 0.0164  | 2184.0259 | [0.000]   |
| 10      | 0.0907 | -0.0016 | 2227.4545 | [0.000]   |
| 11      | 0.0646 | 0.0045  | 2249.4955 | [0.000]   |
| 12      | 0.0693 | 0.0058  | 2274.8809 | [0.000]   |
| 13      | 0.0667 | 0.0271  | 2298.4003 | [0.000]   |
| 14      | 0.0725 | 0.0196  | 2326.1802 | [0.000]   |
| 15      | 0.0625 | 0.0081  | 2346.8292 | [0.000]   |
| 16      | 0.0578 | -0.0002 | 2364.4783 | [0.000]   |
| 17      | 0.0709 | 0.0259  | 2391.0697 | [0.000]   |
| 18      | 0.0659 | 0.0206  | 2414.0172 | [0.000]   |
| 19      | 0.0524 | -0.0031 | 2428.5372 | [0.000]   |
| 20      | 0.0605 | 0.0094  | 2447.9124 | [0.000]   |
| 21      | 0.0468 | -0.0004 | 2459.5151 | [0.000]   |
| 22      | 0.0574 | 0.0187  | 2476.9237 | [0.000]   |
| 23      | 0.0453 | 0.0016  | 2487.7906 | [0.000]   |
| 24      | 0.0567 | 0.0182  | 2504.8288 | [0.000]   |
| 25      | 0.0499 | 0.0034  | 2518.0001 | [0.000]   |
| 26      | 0.0465 | 0.0063  | 2529.4625 | [0.000]   |
| 27      | 0.0616 | 0.0201  | 2549.5455 | [0.000]   |
| 28      | 0.0675 | 0.0323  | 2573.6656 | [0.000]   |
| 29      | 0.0492 | -0.0084 | 2586.4856 | [0.000]   |
| 30      | 0.0455 | -0.0076 | 2597.4527 | [0.000]   |
| 31      | 0.0497 | 0.0072  | 2610.5326 | [0.000]   |
| 32      | 0.0538 | 0.0231  | 2625.886  | [0.000]   |
| 33      | 0.0485 | 0.0077  | 2638.3448 | [0.000]   |
| 34      | 0.0513 | 0.0064  | 2652.2963 | [0.000]   |
| 35      | 0.0797 | 0.0395  | 2685.9376 | [0.000]   |
| 36      | 0.0643 | 0.0134  | 2707.8802 | [0.000]   |
| 37      | 0.0548 | -0.0086 | 2723.8035 | [0.000]   |



Figura 4.5: Principales Estadísticos



En el cuadro 4-3, al igual que en la tabla anterior se muestra la no estacionariedad de la serie de los retornos al cuadrado, al rechazarse la hipótesis nula de estacionariedad de los retornos. En la prueba de Ljung Box , me indica la presencia de correlación serial, debido a que la probabilidad es menor al nivel de significancia del 5 por ciento.

Basados en ambas tablas, me dan señales de la posible necesidad de modelar la media y la varianza condicional del proceso. Por otro lado para examinar la dependencia no lineal , se utilizara la prueba de no linealidad BDS (Brock et.al.,1987) teniendo los resultados en el siguiente cuadro 4-4 .

Cuadro 4.4: Prueba de BDS de retornos porcentuales para distintas dimensiones

Sample: 1/03/1997 3/10/2017 Included observations: 5266

| Dimensión | BDS Statistic | Std. Error | z-Statistic | Prob.  |
|-----------|---------------|------------|-------------|--------|
| 2         | 0.043757      | 0.001537   | 28.47221    | 0.0000 |
| 3         | 0.084327      | 0.002446   | 34.47115    | 0.0000 |
| 4         | 0.115426      | 0.002919   | 39.54204    | 0.0000 |
| 5         | 0.134492      | 0.003049   | 44.10488    | 0.0000 |
| 6         | 0.145334      | 0.002948   | 49.30243    | 0.0000 |

Fuente: Elaboración Propia.

En el cuadro 4-4, se muestra la prueba de dependencia no lineal de BDS(Brock et.al.,1987), las cuales nos muestran la existencia de dependencia no lineal en los retornos, es decir se da un patrón no lineal se rechaza la hipótesis nula de que son variables aleatorias independientes e idénticamente distribuidas reforzando la evidencia encontrada anteriormente, por lo tanto estamos en un entorno de movimiento browniano fraccional.

En la figura 4-6, se presenta el histograma de los retornos acompañado con el estimador no paramétrico de Kernel de la función de densidad de probabilidad de los retornos porcentuales y la función de densidad bajo normalidad, denotando el exceso de curtosis en la serie en comparación con la función de densidad bajo la hipótesis de normalidad, indicando que la función de densidad de los retornos muestran colas pesadas. Esto evidencia que retornos de gran magnitud tanto positivos como negativos se dan con mayor frecuencia en comparación de lo esperado en condiciones de hipótesis de normalidad.

Figura 4.6: Histograma de Densidad de Kernel para retornos porcentuales

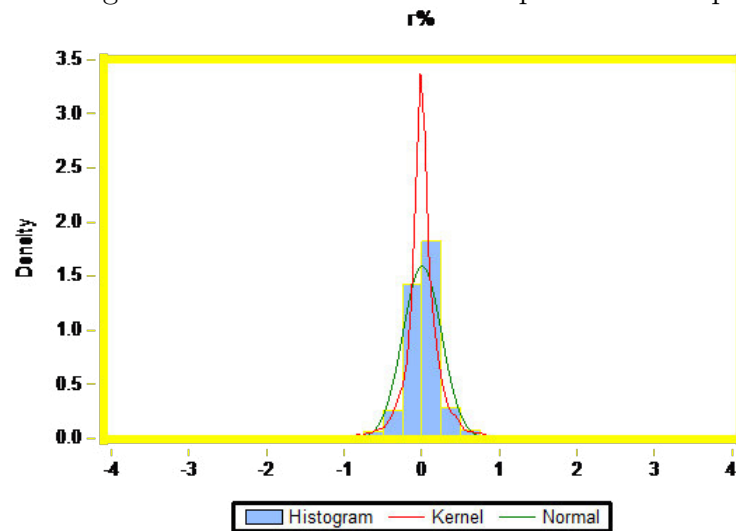
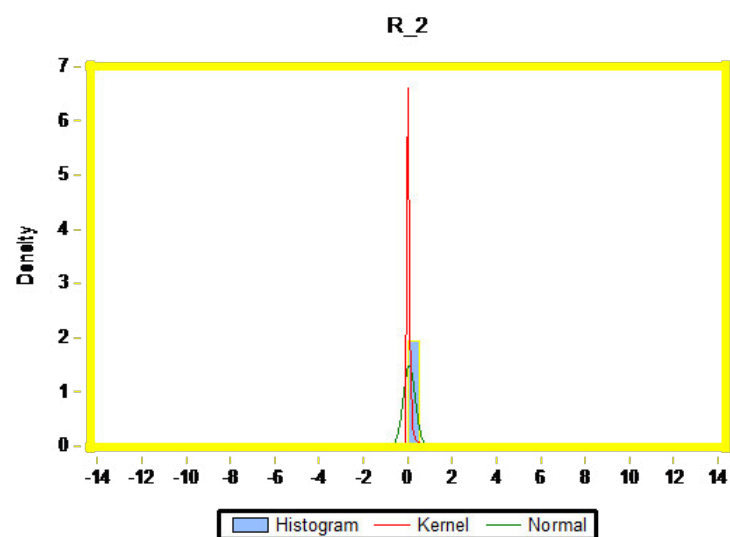


Figura 4.7: Histograma de Densidad de Kernel para retornos porcentuales al cuadrado



En la figura 4-7, se presenta el histograma de los retornos al cuadrado junto con el estimador no paramétrico de Kernel de la función de densidad de probabilidad de los retornos porcentuales y la función de densidad bajo normalidad, denotando el exceso de curtosis en la serie en comparación con la función de densidad bajo la hipótesis de normalidad, indicando que la función de densidad de los retornos muestran colas pesadas. Esto evidencia que retornos de gran magnitud tanto positivos como negativos se dan con mayor frecuencia en comparación de lo esperado en condiciones de hipótesis de normalidad.

Luego de los resultado anteriormente revisados que muestra evidencia empírica de memoria a largo plazo en los retornos porcentuales. Se investigara con mayor profundidad dicha memoria en los retornos porcentuales así como en sus cuadrados y sus valores absolutos, en primer lugar se presentan los correlogramas muestrales para los 36 primeros rezagos en las figuras 4-8,4-9,4-10.

Figura 4.8: Correlogramas para los retornos porecentuales

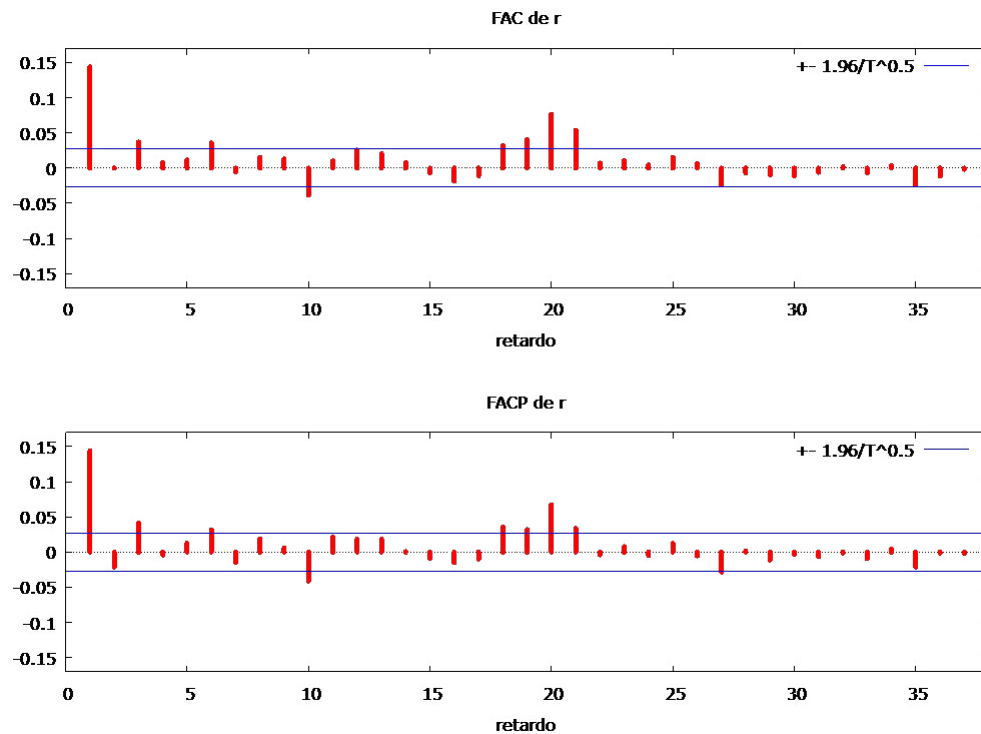


Figura 4.9: Correlogramas para los retornos porcentuales al cuadrado

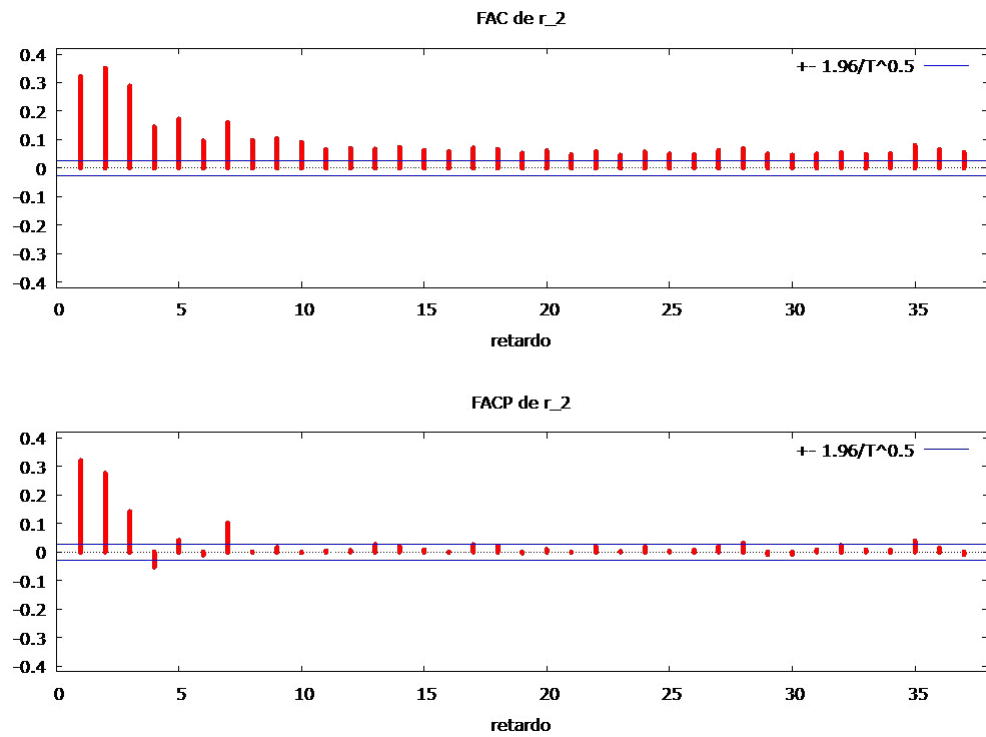
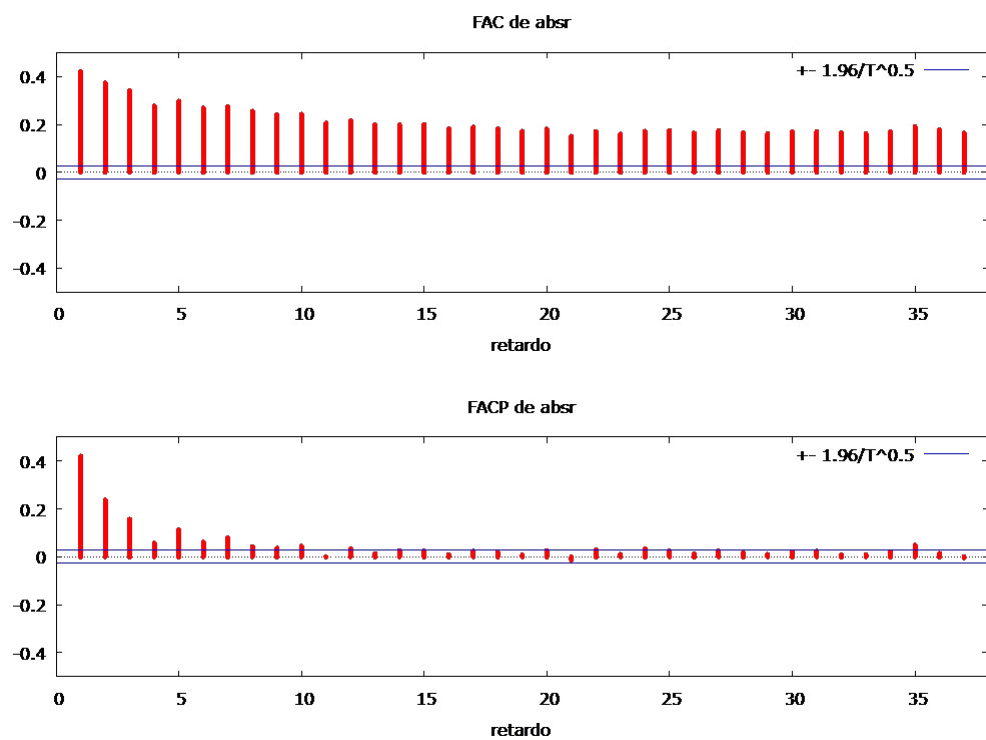


Figura 4.10: Correlogramas para los retornos porcentuales en valor absoluto



Al observar los correlogramas muestrales de las figuras anteriores, se distingue claramente señales de memoria larga (persistencia en la volatilidad) de las series de los retornos al cuadrado y en valor absoluto debido a la presencia de decrecimiento hiperbólico de la autocorrelaciones las cuales sugieren la presencia de persistencia en la volatilidad (memoria a largo plazo). Pero en el correlograma de los retornos porcentuales parece no evidenciar señales de memoria larga.

### 4.2.2. Pruebas de Memoria Larga

En el análisis estadístico del apartado anterior surgió evidencia de la existencia de una posible memoria larga en el proceso de volatilidad, por lo cual se requiere de un conjunto de test que verifiquen dicha hipótesis (memoria larga), teniendo en consideración que las pruebas de memoria larga presentadas a continuación fueron hechas en el programa estadístico oxmetrics y Gretl.

#### Prueba de verosimilitud local de Whittle

Es indispensable para la aplicación de esta prueba de memoria larga utilizar un ancho de banda la cual indica el número de frecuencias utilizadas en la estimación, en este caso nos basamos en el estudio de simulación de Robinson (1995) que sugiere los valores siguientes:

$$m = T^{0,5}, T^{0,73}, T^{0,82}.$$

El estudio cuenta con  $T = 5265$  observaciones, entonces al aplicar con estos anchos de banda, tendría los siguientes resultados:

Cuadro 4.5: Pruebas para retornos porcentuales

| Ancho de banda    | d     | Err. Est. | Razon t | Valor de p |
|-------------------|-------|-----------|---------|------------|
| $T^{0,5} = 73$    | 0.150 | 0.059     | 2.558   | 0.011      |
| $T^{0,64} = 241$  | 0.101 | 0.032     | 3.150   | 0.002      |
| $T^{0,73} = 521$  | 0.062 | 0.022     | 2.853   | 0.004      |
| $T^{0,82} = 1126$ | 0.061 | 0.015     | 4.110   | 0.000      |

Fuente: Elaboración Propia.

Cuadro 4.6: Pruebas para retornos porcentuales al cuadrado

| Ancho de banda    | d      | Err. Est. | Razon t | Valor de p |
|-------------------|--------|-----------|---------|------------|
| $T^{0,5} = 73$    | 0.3579 | 0.0585    | 6.1156  | 0.0000     |
| $T^{0,64} = 241$  | 0.2106 | 0.0322    | 6.5375  | 0.0000     |
| $T^{0,73} = 521$  | 0.2531 | 0.0219    | 11.5557 | 0.0000     |
| $T^{0,82} = 1126$ | 0.3516 | 0.0149    | 23.5967 | 0.0000     |

Fuente: Elaboración Propia.

Cuadro 4.7: Pruebas para retornos porcentuales en valor absoluto

| Ancho de banda    | d      | Err. Est. | Razon t | Valor de p |
|-------------------|--------|-----------|---------|------------|
| $T^{0,5} = 73$    | 0.4781 | 0.0585    | 8.1696  | 0.0000     |
| $T^{0,64} = 241$  | 0.3769 | 0.0322    | 11.7030 | 0.0000     |
| $T^{0,73} = 521$  | 0.3751 | 0.0219    | 17.1252 | 0.0000     |
| $T^{0,82} = 1126$ | 0.3291 | 0.0149    | 22.0859 | 0.0000     |

Fuente: Elaboración Propia.

En los cuadros 4-5,4-6,4-7, se observa la confirmación de la evidencia de existencia de memoria larga en el proceso de volatilidad condicional pues todos los valores de la probabilidad no son estadísticamente significativas , pero también notamos que a diferencia del análisis descriptivo y grafico en la serie de los retornos porcentuales en la cual no se evidenciaba de manera clara la presencia de memoria larga mientras que en la prueba de verosimilitud local de Whittle nos indica de manera más evidente el rechazo de la hipótesis nula de memoria corta, es decir en la tres series el valor de la probabilidad es menor al 5 por ciento como nivel de significancia por lo tanto se rechazaría la hipótesis nula de memoria corta, en consecuencia según la prueba de verosimilitud local de whittle, me indica la existencia de memoria larga. Si consideramos el error estándar de la estimación se desarrolla las pruebas t y en cada una de las tres series es decir en los retornos porcentuales y sus transformaciones el valor de t es mayor a 2 en valor absoluto por lo tanto también se aceptaría la hipótesis alternativa de memoria larga , como conclusión se evidencia memoria larga en el proceso de volatilidad condicional.

### Prueba de Geweke y Porter Hudak

La estimación del parámetro de integración fraccional  $d$ , se basa en el supuesto de que  $E[\log \frac{I_j}{f(w_j)}] = -C$  donde  $I_j$  es el periodograma y  $f(w_j)$  es la función densidad del proceso para todo  $j$  ( $c$  , constante de Euler), (Hurvich and Beltrao(1994)).

Hurvich and Beltrao(1994) refiere que si  $d \neq 0$  , entonces el limite del primer termino de la ecuacion anterior depende de  $j$  lo que genera valores distinto de  $-C$ , si además  $d \neq 0$  y  $j$  es fijo, por lo tanto  $\frac{I_j}{f(w_j)}$  no convergeria en distribución a  $\frac{1}{2}\chi_2^2$  y  $\lim_{j \rightarrow \infty} E[\log \frac{I_j}{f(w_j)}] \neq -C$  (Hurvich and Beltrao, 1993, Robinson, 1995a), entonces se concluye que la distribución limite y la esperanza limite de  $\frac{I_j}{f(w_j)}$  dependen de  $j$ .

Como consecuencia de estas distorsiones en el comportamiento del periodograma en las bajas frecuencias , se realiza una regresión de ecuación sobre ecuación omitiendo un bloque de frecuencias muy bajas para las ordenadas del periodograma.

Por otro lado Robinson(1995a ) menciona que si el objetivo es obtener buenas propiedades asintóticas de la estimación de  $d$ , no es muy importante  $l > 0$ ,  $J$  es un entero mayor o igual a 1. Cuando  $j$  es mayor a 1 a un agrupamiento de las contribuciones desde las frecuencias adyacentes. Geweke and Porter and Hudak, mediante un proceso de simulación muestran como  $M = n^{1/2}$  pudiera ser mas adecuado,sabiendo que  $M$  es el número de frecuencias usadas en la regresión.

Cuadro 4.8: Pruebas para retornos porcentuales

| Ancho de banda | $d$      | Err. Est. | Razon t    | Valor de p |
|----------------|----------|-----------|------------|------------|
| $T^{0.5} = 73$ | 0.154812 | 0.11273   | 1.37329903 | 0.174      |

Fuente: Elaboración Propia.

Según los resultados de las pruebas de Gewek - Porter y Hudak de los cuadros 4-8 y 4-9, Mediante estudios de simulación muestran como  $M = n^{1/2}$  parece ser apropiado, siendo M el número de frecuencias usadas en la regresión. Entonces considerando un ancho de banda  $T^{1/2} = 73$ , los resultados muestran que no hay evidencia suficiente según esta prueba para rechazar la hipótesis nula de memoria corta en los retornos porcentuales al tener un valor de probabilidad 0.174 superior mayor al 5 por ciento de significancia y si tomamos en cuenta el valor de  $t = 1,37329903$  al ser menor a 2 en valor absoluto no es estadísticamente significativa, a diferencia de las series de retornos al cuadrado y valor absoluto en las cuales si encuentran suficiente evidencia para rechazar la hipótesis nula de memoria corta debido a que el valor de la probabilidad es menor al 5 por ciento del nivel de significación, por lo tanto se concluye que existe memoria larga en el proceso de la volatilidad condicional, si tomamos en cuenta que el valor de t tanto en la serie del cuadrado de los retornos como en el valor absoluto son mayores a 2 por lo tanto confirma aceptar la hipótesis alternativa de memoria larga.

Cuadro 4.9: Pruebas para retornos porcentuales al cuadrado

| Ancho de banda | d        | Err. Est. | Razon t    | Valor de p |
|----------------|----------|-----------|------------|------------|
| $T^{0,5} = 73$ | 0.390014 | 0.0752275 | 5.18446047 | 0          |

Fuente: Elaboración Propia.



## Prueba de Coeficiente de Hurts

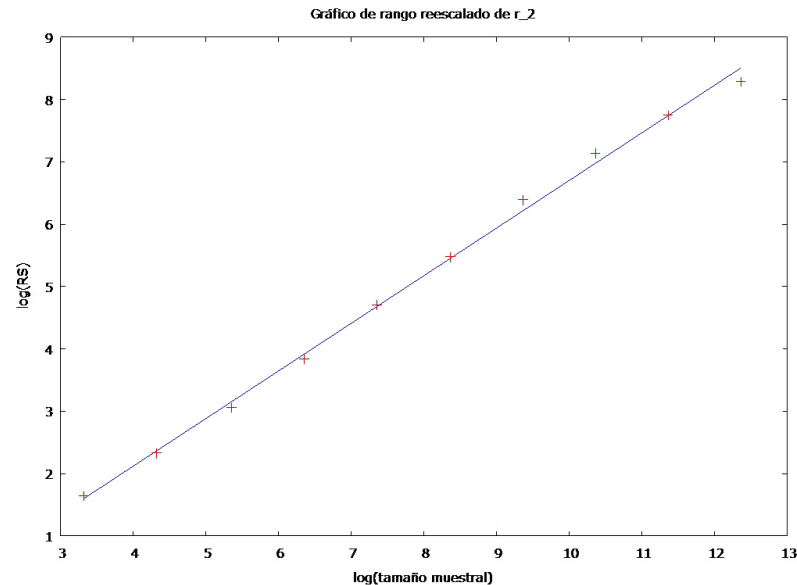
A continuación en el cuadro 4-10, se muestra el proceso realizado para obtener el coeficiente de hurst, dividiendo la muestra en 1, 2, 3, hasta 50 partes para luego obtener el estadístico R/S. Con estos datos obtenemos el coeficiente de hurst que representa la pendiente de la recta del gráfico de la figura 5-11 es decir el ajuste de la regresión. De dicho ajuste en el cuadro 4-12 se muestra el coeficiente de hurts de la serie del cuadrado de los retornos el cual me indica un valor de 0.76409 siendo mayor a 0.5 y menor a 1 entonces indicaria la sospecha de un movimiento browniano fraccionario con un parámetro de memoria larga  $d > 0$  por lo tanto la existencia de dependencia de largo plazo en los retornos porcentuales al cuadrado del tipo de cambio, es decir memoria larga en el proceso de la varianza condicionada, se consideraría dicha serie como caótica muy sensible a cambios en las condiciones iniciales, es decir, la tendencia de la serie seria la misma al periodo anterior. Se tendrá un 76.5 por ciento de probabilidad de que la tendencia sea la misma del periodo anterior, por otro lado dicho coeficiente de hurts me permite calcular la dimensión fractal propiedad característica de una estructura fractal, al definirse la dimensión como  $D = 2 - H$  dicho resultado será un numero fraccionario cercano a 1.235905. La interpretación esta relacionada con la complejidad del sistema, es decir el número entero próximo a la dimensión fractal nos indica el mínimo número de variables que se necesita para modelar el sistema y colocar un límite inferior al número de grados de libertad.

Cuadro 4.10: Cifras de rango reescalado del cuadrado de los retornos

| R.2 | Tamaño | RS(medio) | log(Tamaño) | log(RS) |
|-----|--------|-----------|-------------|---------|
| 1   | 5265   | 312.81    | 12.362      | 8.2891  |
| 2   | 2632   | 215.07    | 11.362      | 7.7487  |
| 3   | 1316   | 141.04    | 10.362      | 7.1399  |
| 4   | 658    | 83.988    | 9.3619      | 6.3921  |
| 5   | 329    | 44.534    | 8.3619      | 5.4768  |
| 10  | 164    | 26.049    | 7.3576      | 4.7032  |
| 20  | 82     | 14.305    | 6.3576      | 3.8384  |
| 30  | 41     | 8.3145    | 5.3576      | 3.0556  |
| 40  | 20     | 5.0265    | 4.3219      | 2.3296  |
| 50  | 10     | 3.1253    | 3.3219      | 1.644   |

Fuente: Elaboración Propia.

Figura 4.11: Ajuste del logaritmo de reescalamiento de la serie del cuadrado de los retornos

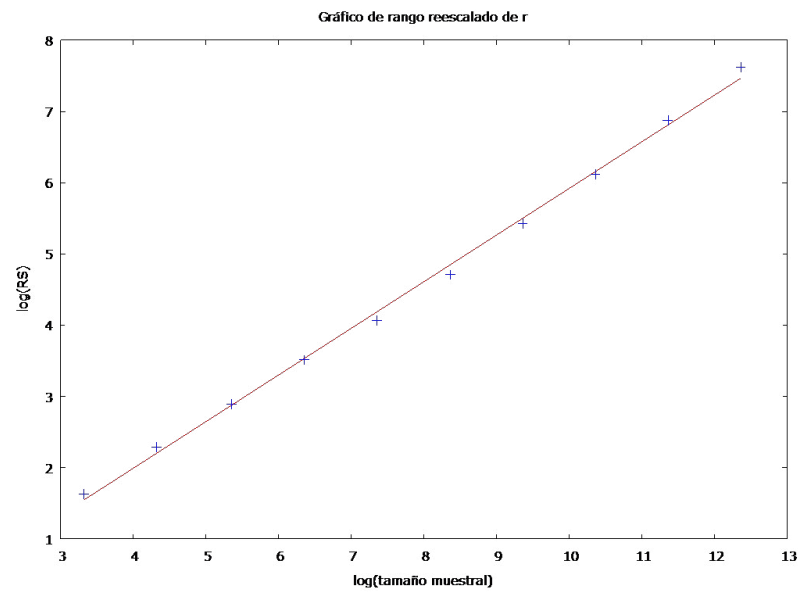


Luego en el cuadro 4-11, se muestra el proceso realizado para obtener el coeficiente de hurst para la serie de los retornos porcentuales del tipo de cambio, dividiendo la muestra en 1, 2, 3, hasta 50 partes para luego obtener el estadístico R/S. Con estos datos obtenemos el coeficiente de hurst que representa la pendiente de la recta del gráfico de la figura 4-10 es decir el ajuste de la regresión. El cálculo del exponente de Hurts del cuadro 4-12, indica que el valor es de 0.655 (persistencia) mayor a 0.5 y menor a 1, señala una posible persistencia en la serie por lo tanto un movimiento browniano fraccionario (la existencia de memoria larga) se identificaría como una serie caótica muy sensible a cambios en las condiciones iniciales, entonces se tendrá un 65.5 por ciento de probabilidad de que en el periodo siguiente se genere la misma tendencia del periodo anterior y como consecuencia dicho coeficiente de hurts permite calcular la dimensión fractal propiedad característica de una estructura fractal, al definirse la dimensión como  $D = 2 - H$  dicho resultado será un numero fraccionario cercano a 1.345037. La interpretación esta relacionada con la complejidad del sistema, es decir el número entero próximo a la dimensión fractal nos indica el mínimo número de variables que se necesita para modelar el sistema y colocar un límite inferior al número de grados de libertad.

Cuadro 4.11: Cifras de rango reescalado de r

| R  | Tamaño | RS(medio) | log(Tamaño) | log(RS) |
|----|--------|-----------|-------------|---------|
| 1  | 5265   | 196.77    | 12.362      | 7.6204  |
| 2  | 2632   | 117.43    | 11.362      | 6.8756  |
| 3  | 1316   | 69.431    | 10.362      | 6.1175  |
| 4  | 658    | 43.084    | 9.3619      | 5.4291  |
| 5  | 329    | 26.153    | 8.3619      | 4.7089  |
| 10 | 164    | 16.74     | 7.3576      | 4.0652  |
| 20 | 82     | 11.405    | 6.3576      | 3.5116  |
| 30 | 41     | 7.4076    | 5.3576      | 2.889   |
| 40 | 20     | 4.8877    | 4.3219      | 2.2892  |
| 50 | 10     | 3.102     | 3.3219      | 1.6332  |

Figura 4.12: Ajuste del logaritmo de reescalamiento de la serie r



Cuadro 4.12: Coeficiente de Hurts

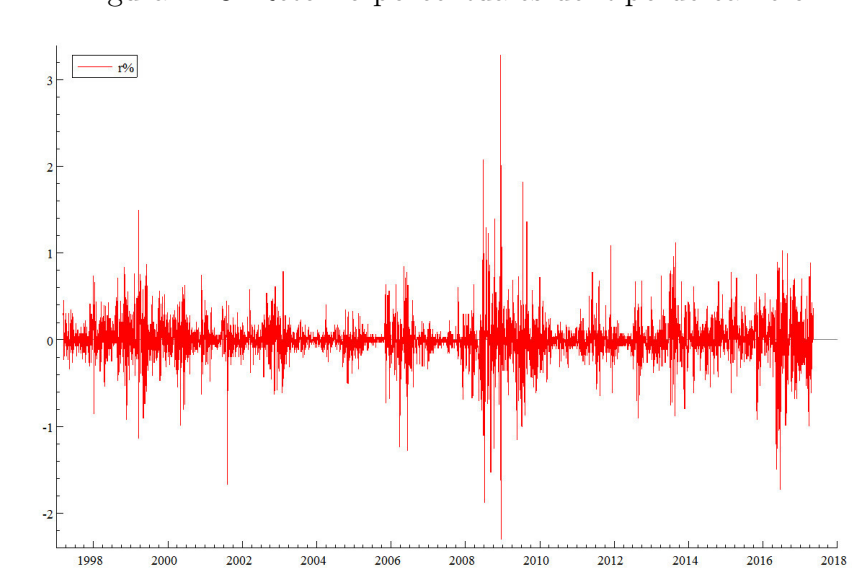
| Serie                    | Exponente de Hurst estimado |
|--------------------------|-----------------------------|
| cuadrado de los retornos | 0.764095                    |
| retornos porcentuales    | 0.654963                    |

Estos resultados demuestran que los retornos porcentuales y sus transformaciones de la serie de tipo de cambio, no siguen un comportamiento aleatorio puro, random walk, sino que están sesgados, entonces decimos que siguen lo que se llama un comportamiento browniano fraccional. Si nos damos cuenta se denota una mayor evidencia de persistencia en la serie del cuadrado de los retornos con un coeficiente de hurts de 0.764095 mayor al 0.654963 coeficiente de hurts de los retornos, los cuales son confirmados por las pruebas de verosimilitud local de Whittle y de GPH y el análisis estadístico preliminar, entonces podríamos concluir la presencia de memoria larga en el proceso de volatilidad condicional, la cual se modelara mediante un FIGARCH.

### 4.3. Resultados Metodológicos

En esta sección se realizará la estimación del modelo de memoria larga FIGARCH es decir un proceso estocástico gaussiano de tipo movimiento browniano fraccional con parámetro de memoria larga  $d$  de una serie de tiempo obtenido del Banco central de reserva del Perú de tipo de cambio diario nominal dólar-sol desde el 3 de Enero de 1997 hasta el 9 de Marzo del 2017, generando un total de 5625 observaciones. Como referencia, en la figura 4.13 se representa los retornos porcentuales del tipo de cambio bajo consideración, la cual indica la existencia de la volatilidad de la serie en el periodo de análisis, del mismo modo la función de auto correlación del cuadrado de los retornos indica una caída hiperbólica como se muestra en la figura 4.14, indicando la posibilidad de modelar la serie con un modelo de memoria larga cuyas pruebas ya se indicaron en el apartado anterior.

Figura 4.13: Retorno porcentuales del tipo de cambio



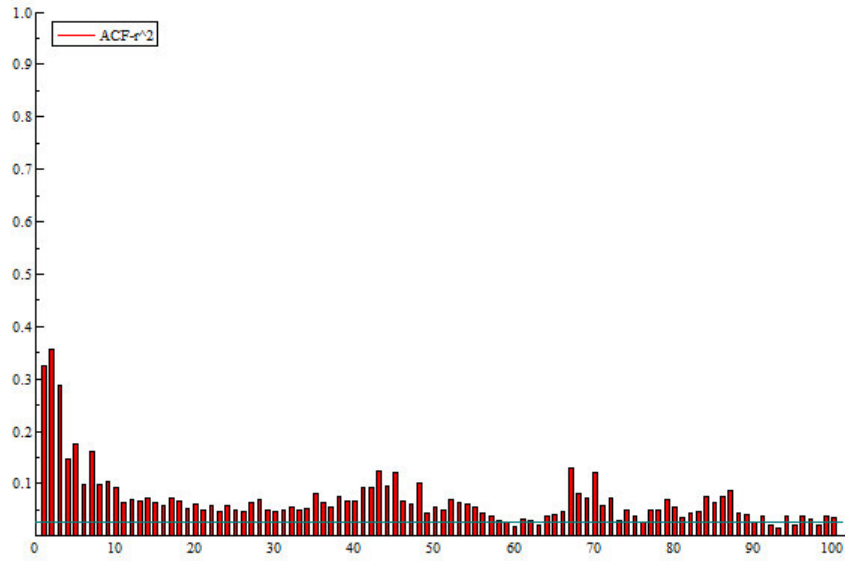
La estimación del modelo FIGARCH para la data existente es la siguiente:

$$r_t = 100 \log(s_t/s_{t-1}) = \mu + \epsilon_t, \epsilon_t h_t^{-1/2} \sim N(0, 1)$$

$$h_t = \alpha_0 + \beta_1 h_{t-1} + [1 - \beta_1 L - (1 - \phi_1 L)(1 - L)^d] \epsilon_t^2, t = 1, 2, \dots, 5265$$

En este caso se aplicará la estructura teórica del movimiento browniano fraccional mediante la modelación de los retornos porcentuales logarítmicos  $r_t$  los cuales tienen incrementos dependientes a largo plazo definidos en la primera parte de la tesis.

Figura 4.14: Funcion de autocorrelación del cuadrado de los retornos



La estimación es mediante la QMLE bajo la suposición de la normalidad condicional, en el cuadro 4.13 entre paréntesis se indican los errores estándar robustos de cada estimación en dicho cuadro se indican los modelos estimados de figarch, los cuantiles  $Q(20)$  y  $Q(20)^2$  se refieren a la prueba de Ljung Box Portmanteau (Apendice 2) de la correlación serial de hasta el orden 20 en los residuos estandarizados y los residuos estandarizados al cuadrado respectivamente, en cada columna se describen las estimaciones de cada modelo, en la primera columna se describe al modelo  $(p, d, q) = (0, 0, 0)$  la cual corresponde al modelo GARCH(0,0), en la segunda y tercera columna se especifican diferentes modelos FIGARCH. En las cuales se puede notar que el valor de logaritmo de verosimilitud de FIGARCH(1,d,1) es más alto dicho valor me indica entonces la idoneidad del modelo FIGARCH(1,d,1).

Por otro lado si usamos los criterios de información dados en la tabla 4.14, el modelo FIGARCH(1,0.65,1), es el proceso que mejor se adapta a la data en análisis de tipo de cambio debido al menor valor comparando entre todos los criterios de información de cada uno de los modelos estimados indicados en la tabla 4.14, por lo tanto se fortalece la propuesta del modelo FIGARCH(1,0.65,1) como el modelo más idóneo para el conjunto de datos existente.

Si analizamos el cuadro 4.15, se señalan dos de las diferentes medidas que se utilizan para comparar la precisión entre modelos de pronóstico MSE (error

cuadrático medio) y RSME(raíz del error cuadrático medio) que corresponde al pronóstico de la varianza condicional. Se nota que el modelo FIGARCH(1,0.65,1) da los mejores pronósticos representados en mínimos valores en MSE Y RMSE con valores de 0.0001511 0.01229 respectivamente.

Cuadro 4.13: Estimación de modelos FIGARCH

| (p,d,q)             | (0,0,0)               | (1,d,0)                | (1,d,1)                |
|---------------------|-----------------------|------------------------|------------------------|
| $\mu$               | 0.001678<br>0.0036107 | -0.008167<br>0.0018807 | -0.007849<br>0.0018779 |
| $\alpha_0$          | 0.058414              | 0.000133               | 0.000165               |
| $\beta_1$           | -                     | 0.262035               | 0.647364<br>0.075995   |
| $\phi_1$            | -                     | -                      | 0.293989<br>0.073052   |
| d                   | -                     | 0.527594<br>0.05746    | 0.649139<br>0.07466    |
| S                   | 0.65719               | 0.18541                | 0.23524                |
| K                   | 18.943                | 5.3988                 | 5.5193                 |
| Q( 20)              | 148.662               | 145.449                | 135.622                |
| Q( 20) <sup>2</sup> | 2069.97               | 17.6941                | 14.4002                |
| LogLikelihood:      | 5.005                 | 1697.577               | 1714.682               |

Fuente: Elaboración Propia.

Cuadro 4.14: Criterio de información para la elección del modelo

| (p,d,q)    | Akaike    | Schwarz   | Shibata   | Hannan-Quinn |
|------------|-----------|-----------|-----------|--------------|
| (0,0,0)    | -0.001341 | 0.001519  | -0.001342 | -0.000333    |
| (1,0.53,0) | -0.755892 | -0.750173 | -0.755894 | -0.753877    |
| (1,0.65,1) | -0.763081 | -0.755931 | -0.763083 | -0.760561    |

Fuente: Elaboración Propia.

Si consideramos una medida de selección entre modelos de pronóstico adicional, se tiene en la tabla 4.15, el coeficiente de desigualdad de Theil la cual me permite analizar la efectividad del modelo de predicción seleccionado cuyo valor puede ser mayor a 1, cuanto más cercano a cero mejor es la efectividad del modelo, es decir entre los tres modelos estimados se elegirá el más cercano a cero, en el cuadro 5.3 los modelos FIGARCH(1,0.53,0) con un valor de 0.345 y el modelo FIGARCH(1,0.65,1) con valor de 0.346 son casi de la misma efectividad, pero se podría concluir en función a los análisis anteriormente realizados que a pesar de ser marginalmente mayor coeficiente de Theil pero en vista de la evidencia de las anteriores medidas me indicarían, que el mejor modelo seria el proceso browniano fraccional de FIGARCH(1,0.65,1) o de memoria larga.

Cuadro 4.15: Medidas de error de los pronósticos de la varianza condicional

| (p,d,q)    | MSE       | RMSE    | TIC    |
|------------|-----------|---------|--------|
| (0,0,0)    | 0.002162  | 0.0465  | 0.614  |
| (1,0.53,0) | 0.0001604 | 0.01266 | 0.3449 |
| (1,0.65,1) | 0.0001511 | 0.01229 | 0.3457 |

Fuente: Elaboración Propia.

Cuadro 4.16: Q-Statistics on Squared Standardized Residuals Ljung-Box

|         |         |             |
|---------|---------|-------------|
| Q(10) = | 10.9355 | [0.2053786] |
| Q(20) = | 14.4002 | [0.7026523] |
| Q(50) = | 35.917  | [0.9006973] |

Fuente: Elaboración Propia.

En el siguiente cuadro 4-16, se muestra la prueba de Ljung- Box de los residuales estandarizados al cuadrado en donde me indica que se dan señales de que el modelo propuesto para el proceso de volatilidad captura adecuadamente la estructura de correlación debido a que la probabilidad asociada es mayor al 0,05 aceptándose la hipótesis nula de ausencia de autocorrelación.

En el siguiente cuadro 4-17, se muestra la prueba de Engle (1982) de los residuales estandarizados al cuadrado en la cual se acepta la hipótesis nula de homocedasticidad en los residuos debido a que la probabilidad asociada al estadístico nR2 es mayor a 0.05 entonces confirma la no existencia de posible correlacion serial significativa en los residuales.

De acuerdo con Hsieh (1989) la varianza condicional se encuentra bien especificada cuando el exceso de curtosis de los residuos estandarizados no exceden al exceso de curtosis de los residuos no estandarizados. Entonces para este caso el exceso de curtosis de los residuos estandarizados es 4.5 y el exceso de curtosis de los residuales no estandarizados es 18.9 por lo tanto el modelo está muy bien especificado.

Cuadro 4.17: ARCH test on Squared Standardized Residuals

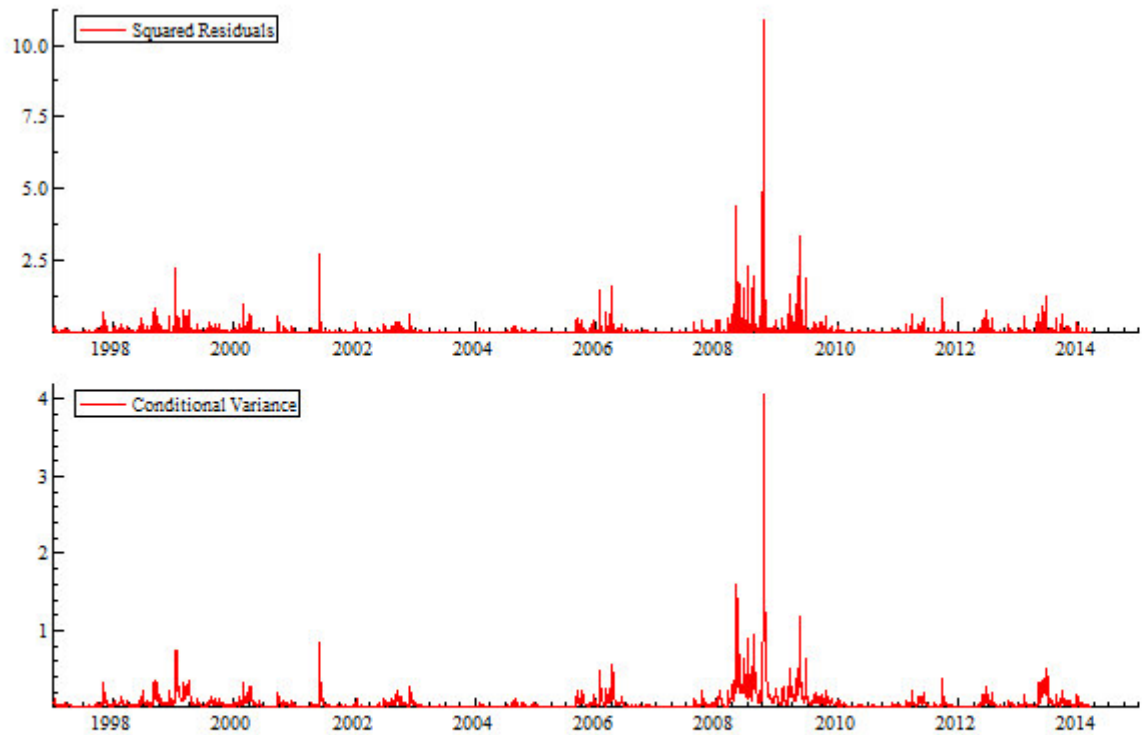
|            |              |         |          |
|------------|--------------|---------|----------|
| ARCH test: | F(2,4474) =  | 0.10625 | [0.8992] |
| ARCH test: | F(5,4468) =  | 1.7345  | [0.1231] |
| ARCH test: | F(10,4458) = | 1.052   | [0.3964] |

Fuente: Elaboración Propia.

Al observar el cuadro 4.18 el contraste de Box - Pierce se evidencia con un nivel de significancia dado que la serie generada de residuos no se encuentra correlacionada. Por lo cual se considera que el modelo propuesto captura de mejor

manera la estructura de correlación del proceso de los retornos en volatilidad a través de un FIGARCH(1,0.65,1). Por lo tanto la varianza condicionada estimada seria la de la figura 4.15.

Figura 4.15: La varianza condicional estimada





Cuadro 4.18: Correlograma para Std.Residuales al cuadrado utilizando desviación típica  $1/T^{0,5}$

| RETARDO | FAC     |    | FACP    |    | Estad-Q. | [valor p] |
|---------|---------|----|---------|----|----------|-----------|
| 1       | 0       |    | 0       |    | 0        | [0.999]   |
| 2       | 0.0069  |    | 0.0069  |    | 0.213    | [0.899]   |
| 3       | 0.0177  |    | 0.0177  |    | 1.6144   | [0.656]   |
| 4       | -0.0185 |    | -0.0185 |    | 3.1437   | [0.534]   |
| 5       | 0.0354  | ** | 0.0351  | ** | 8.7557   | [0.119]   |
| 6       | -0.0051 |    | -0.0052 |    | 8.8727   | [0.181]   |
| 7       | -0.015  |    | -0.0149 |    | 9.8874   | [0.195]   |
| 8       | 0.0124  |    | 0.0109  |    | 10.5737  | [0.227]   |
| 9       | -0.0055 |    | -0.0039 |    | 10.7102  | [0.296]   |
| 10      | 0.0071  |    | 0.006   |    | 10.9355  | [0.363]   |
| 11      | -0.0078 |    | -0.0083 |    | 11.2083  | [0.426]   |
| 12      | -0.0139 |    | -0.0124 |    | 12.0821  | [0.439]   |
| 13      | -0.0039 |    | -0.0052 |    | 12.1517  | [0.515]   |
| 14      | -0.0158 |    | -0.0149 |    | 13.2745  | [0.505]   |
| 15      | -0.0004 |    | -0.0003 |    | 13.2752  | [0.581]   |
| 16      | -0.008  |    | -0.0078 |    | 13.5648  | [0.631]   |
| 17      | -0.0077 |    | -0.0062 |    | 13.8307  | [0.679]   |
| 18      | 0.0091  |    | 0.0085  |    | 14.2035  | [0.716]   |
| 19      | 0.0018  |    | 0.003   |    | 14.2177  | [0.771]   |
| 20      | -0.0064 |    | -0.0066 |    | 14.4002  | [0.810]   |
| 21      | -0.0238 |    | -0.0243 |    | 16.9553  | [0.714]   |
| 22      | 0.0109  |    | 0.0121  |    | 17.4873  | [0.736]   |
| 23      | -0.0072 |    | -0.0078 |    | 17.7195  | [0.773]   |
| 24      | 0.0047  |    | 0.005   |    | 17.8196  | [0.812]   |
| 25      | 0.0244  |    | 0.0238  |    | 20.503   | [0.720]   |
| 26      | -0.0075 |    | -0.0059 |    | 20.7579  | [0.754]   |
| 27      | -0.0054 |    | -0.0076 |    | 20.8914  | [0.791]   |
| 28      | 0.0034  |    | 0.0023  |    | 20.9449  | [0.828]   |
| 29      | 0.0113  |    | 0.0128  |    | 21.5246  | [0.839]   |
| 30      | 0.008   |    | 0.0057  |    | 21.8134  | [0.861]   |
| 31      | 0.0088  |    | 0.0096  |    | 22.1651  | [0.878]   |
| 32      | -0.0016 |    | -0.0015 |    | 22.1763  | [0.903]   |
| 33      | 0.022   |    | 0.0208  |    | 24.3683  | [0.862]   |
| 34      | 0.0075  |    | 0.0066  |    | 24.6251  | [0.881]   |
| 35      | 0.0043  |    | 0.0033  |    | 24.7089  | [0.902]   |
| 36      | -0.0119 |    | -0.0125 |    | 25.3479  | [0.907]   |

# Capítulo 5

## Conclusiones

### 5.1. Conclusiones

1. Al haberse estudiado la serie de tiempo del tipo de cambio nominal en Perú se ha logrado cumplir con el objetivo de determinar la durabilidad de los shocks en la volatilidad de la serie tipo de cambio nominal de Perú del periodo(1997-2017) comprobando la persistencia de los shocks que tiene una durabilidad a largo plazo al evidenciar empíricamente un orden de integración fraccional de  $d = 0,6491$  en su volatilidad.
2. Como consecuencia de la estimación del parámetro de memoria larga  $d > 0$  se identifica al proceso de varianza condicional como un proceso autosimilar con  $H > 0$  con un coeficiente de Hurts  $H = 0,764095$  en los retornos cuadráticos  $H \neq \frac{1}{2}$  evidenciando una mayor persistencia en la volatilidad de esta manera se concluye rechazando la naturaleza del proceso como un movimiento browniano identificando de este modo al proceso como un movimiento browniano fraccional no estacionario con incrementos no estacionarios debido a que el parámetro de memoria larga  $d = 0,6491$  se encuentra fuera del rango  $(-0,5, 0,5)$  dicho rango garantiza la estacionariedad de sus incrementos y con incrementos dependientes como se evidencia mediante la prueba BDS de los retornos porcentuales señalado en el cuadro 4.4 cuya dependencia es no lineal. De esta manera se logra el objetivo de identificar al proceso de varianza condicional como un movimiento browniano fraccional(sistema caótico).
3. Se logró la estimación y la especificación del modelo Autorregresivo con Heterocedasticidad condicionada Generalizada Fraccionalmente integrado del tipo de cambio nominal del Perú - FIGARCH(1,d,1) señalado en el cuadro 4.13.

$$h_t = 0,000165 + 0,647364h_{t-1} + [1 - 0,647364L - (1 - 0,293989L)(1 - L)^{0,649139}] \epsilon_t^2$$

,  $t = 1, 2, \dots, 5265$

El modelo FIGARCH(1,0.65,1) cumple con todos los test de significancia de sus parámetros y las desigualdades dadas por Baillie(2001) descritas en la sección 3.1.5 para garantizar la no negatividad de la varianza condicionada.

Confirmando de esta manera el hecho de que el modelo propuesto es adecuado respaldado por la prueba de Ljung-Box de los residuos estandarizados al cuadrado en donde me indica de que el modelo propuesto para la volatilidad captura adecuadamente la estructura de correlación evidenciando ausencia de correlación como se muestra en el cuadro 4.16. Del mismo modo el cuadro 4.17 me confirma la ausencia de correlación serial en los residuos a través de la prueba ARCH.

Según el criterio de Hsieh(1989) la varianza condicionada se encuentra bien especificada debido a que el exceso de curtosis de los residuos estandarizados con un valor de 4.5 no excede al exceso de curtosis de los residuos no estandarizados con un valor de 18.9.

4. Se logró desarrollar el modelo FIGARCH(m,d,q) a partir de un modelo sin memoria como un ARCH(p) describiendo sus características y propiedades fundamentales descritos en la metodología econométrica de la sección 3.1.2.
5. En relación a los modelos de volatilidad estimados se muestran buenos resultados en el modelo FIGARCH(1,0.53,1) y aceptables en el modelo GARCH(0,0), sin embargo debido a que en el modelo FIGARCH(1,0.53,1) el valor de logaritmo de verosimilitud es menor que en modelo FIGARCH(1,0.65,1) y en el modelo GARCH(0,0) el valor de logaritmo de verosimilitud es significativamente de menor que en modelo FIGARCH(1,0.65,1) se demuestra la idoneidad del modelo propuesto FIGARCH(1,0.65,1) como se indica en el cuadro 4.13. Por otro lado si nos basamos en los criterios de información del cuadro 4.14 el modelo FIGARCH(1,0.65,1) es el proceso que mejor se adapta a la data en análisis del tipo de cambio. Del mismo modo respalda la idoneidad del modelo propuesto FIGARCH(1,0.65,1) los errores cuadrático medio y raíz del error cuadrático medio evidenciando mejores pronósticos debido a sus mínimos valores como se muestra en el cuadro 4.15. Es importante mencionar que solamente se realizó la comparación de modelos FIGARCH y GARCH debido al objetivo de la tesis específicamente de memoria larga en volatilidad FIGARCH(1,d,1).
6. Mediante la función impulso respuesta acumulativa se cuantificó y demostró la durabilidad del shock en la volatilidad de la varianza condicionada que tiende a cero pero hiperbólicamente a una tasa de  $n^{0.649139}$ . Por otro lado se utilizó la función hipergeométrica para explicar el ruido blanco fraccionario  $F(d-1, 1, 1; 1) = 0$  y  $\lim_{s \rightarrow \infty} \sum_{i=0}^s \varphi(i) = 0$ . Cabe mencionar que debido al parámetro de integracional fraccional estimado  $d = 0,6491$  la durabilidad del shock en la volatilidad desaparece de manera muy lenta, es decir un proceso no estacionario con incrementos no estacionarios y dependientes cuya dependencia es no lineal.

## 5.2. Sugerencias para futuras investigaciones

1. Es importante mencionar que existen investigaciones como la de Sebastián Laurent y Michael Baine(2000) los cuales indican que la posible existencia de memoria larga en la serie es debido a un posible quiebre estructural, encontrando evidencia que sustenta la hipótesis de que la existencia de un posible quiebre estructural generaría un incremento en la memoria de la serie, debido a la extensión que demandaría incluir dicho análisis solo mencionó este fenómeno para futuras investigaciones.
2. Con respecto a la relación quiebre estructural y memoria larga Baillie y Morata(2009)desarrollan un nuevo proceso de memoria larga que incluye la dependencia a largo plazo y quiebre estructural en el proceso de volatilidad denominado A-FIGARCH(p,d,q,k) (Adaptive FIGARCH) dicho modelo constituye un componente estocásticos de memoria larga y otro componente determinístico de proceso de quiebre, dicho tema es sujeto en la actualidad de investigaciones en Nash Boutahr y Travbelsi(2010) se presenta mucha mas información de estructuras de A-FIGARCH Y derivados como FITVGARCH.

# Bibliografía

- [1] Barry R.J.(1981) *Probabilidade*. Instituto de Matematica Pura y Aplicada.
- [2] Billingsley, P.(1968). *Convergence of Probability Measures*. John Wiley Sons New York.
- [3] Brockwell P.J. Davis R.A.Springer Verlag (1987). *Time Series Theory and Methods*. Paris: Mason.
- [4] Brockwell P.J. Davis R.A. (2002). *Introduction to Time Series and Forecasting*. Springer Verlag.
- [5] Cavanzo, A. (2004). *El movimiento browniano fraccional como limite de ciertos tipos de procesos estocásticos*.
- [6] C. Tudor (2002). *Procesos Estocasticos..Aportaciones Matematicas de la Sociedad Matematica Mexicana*.
- [7] C. Fox. (1928). *The asymptotic expansion of generalized hypergeometric functions*. Proc. London Math. Soc. London, 2(27), 389 - 400.
- [8] E. Castano, S. Gallón, and K. Gómez(2008). *Una prueba para el parámetro de diferenciación fraccional..Revista Colombiana de Estadística*, 31:67 - 84, 2008.
- [9] Engle, R. (1982) *Autoregressive conditional heterocedasticity with estimates of the variance of United Kingdom inflations*. *Econometrica* . (50), 987-1007.
- [10] Feller, W. (1971). *An Introduction to Probability Theory and its Applications* . John Willey Sons, New York.
- [11] Gonzales and Nieto. (2008). *Distribución de la estadística de Jarque Bera para la prueba de normalidad en una serie temporal estacionaria con datos faltantes* . Entre Ciencia e Ingenieria, 4:99-114,2008.
- [12] Granger, C. W. J. (1980). *Long memory relationships and the aggregation of dynamic models*. Journal of Econometrics, pp. 227-238.
- [13] Hurvich, C. M., R. Deo and J. Brodsky (1998). *The mean squared error of Geweke and Porter-Hudaks estimator of the memory parameter*.
- [14] Kooperberg, C., and Stone, C. (1991). *A Study of Logspline Density Estimation*. Computational Statistics and Data Analysis 12, 327 - 347.

- [15] Lin, S. (1995). *Stochastic analysis of fractional brownian motion*. Stochastic and Stochastic Reports 55(1-2).
- [16] Lindsay B.(1989). *Moment Matrices: Applications in Mixtures*. The Annals of Statistics 17 No. 2, 722-740.
- [17] Lombardi, M. J. and G. M. Gallo (2005). *Processi a memoria lunga e fra-  
tionally integrated garch*. Quaderni/2005-2, Dipartimento de Statistica Giu-  
seppe Parenti Università degli Studi di Firenze.
- [18] Luengas Diego Dominguez, Esperanza Ardila Romero, John Freddy Moreno  
Trujillo(2010). *Metodología e interpretación del coeficiente de hurts*. Univer-  
sidad Externado de Colombia.
- [19] Maryam Tayefi and T.V.Ramanathan (2012). *An Overview on FIGARCH  
and related Time series models*. Department of Statistics And center for  
Advance Studies, University of Pune , India
- [20] Matematica Azucena Campillo Navarro (2013). *Estudio del Movimiento  
Browniano Fraccionario*. Universidad Autonoma Metropolitana, Mexico.
- [21] Nualart D. (2005). *Fractional Brownian Motion: Stochastic Calculus and  
Applications*. XIV Escuela Latinoamericana de Matemática, Balneario Solís,  
Uruguay .
- [22] Ojeda Echeverri César Andrés(2012). *Una prueba de la eficiencia débil en el  
mercado accionario colombiano*. Universidad Nacional de Colombia.
- [23] Robinson, P.M. (1994a). *Semiparametric analysis of long-memory time se-  
ries*. Annals of Statistics 22, 515-539.
- [24] Robinson, P.M. (1994b). *Time series with strong dependence*. in: C.A. Sims,  
ed., Advances in econometrics - Sixth world congress of the econometric  
Society (Cambridge University Press, Cambridge).
- [25] Rocha E. (2014). *Consistencia de los estimadores de máxima verosimili-  
tud gaussiana y de mínimos cuadrados en procesos autorregresivos de media  
móvil*. Monografía de Licenciatura, Facultad de Ciencias .
- [26] Samorodnitsky G. (2006). *Long Range Dependence*. Foundations and trends  
in stochastic systems.
- [27] Taqqu M.S., Teverovsky V., Willinger W.((1995)). *Estimators for Long-  
Range Dependence: An Empirical Study, Fractals*.3, 785-798.
- [28] Taqqu M. S.(2013). *Benoit Mandelbrot and Fractional Brownian Mo-  
tion* Statistical Science, 28, No. 1, pp. 131-134.
- [29] Taylor, S.(1986). *Modelling financial time series*.(Wiley, Chichester).
- [30] Taylor, S. (1994 ). *Modelling stochastic volatility*. Mathematical Finance 4,  
183-204.

- [31] Tsehernig, R.(1995 ). *Long memory in foreign exchange rates revisited*.Journal of International Financial Markets, Institutions, and Money 5, 53-78. Mathematical Finance 4, 183-204.
- [32] Weiss, A.A.(1986). *Asymptotic theory for ARCH models: Estimation and testing*, *Econometric Theory* 2. 107-131.
- [33] Yajima, Y.(1988). *On estimation of a regression model with long-memory stationary errors*.Journal of International Financial Markets, Annals of Statistics 16, 791-807.

# Apéndices

## 5.3. Apéndice 1: BDS

### 5.3.1. Prueba de no linealidad (BDS)

Tomando la literatura (Cesar Andres Ojeda Echeverri, 2012). Esta prueba supone como en el caso lineal que los errores están independientes e idénticamente distribuidos con media cero y varianza 1 es decir ruido blanco se podrían considerar un movimiento browniano ordinario, siendo el objetivo encontrar una dependencia no lineal en los errores. Entonces se podría denotar como  $x_t = f(\epsilon_{t-1}, \dots, \epsilon_{t-k}) + \epsilon_t g(\epsilon_{t-1}, \dots, \epsilon_{t-k})$  Donde:  $f(\cdot)$  es la media de  $x_t$  y  $g(\cdot)$  es la varianza de  $x_t$ . En el año 1989 Hsieh nos da a conocer un estadístico basado en el tercer momento :

$$\varphi(i, j) \equiv \frac{E[x_t x_{t-i} x_{t-j}]}{E[x_t^3]^{3/2}}$$

Entonces si  $\varphi(i, j) = 0 \forall i, j > 0$ , para todos i.i.d. y sugiere estimarlo :

$$\hat{k}(i, j) = \frac{T^{-1} \sum_t x_t x_{t-i} x_{t-j}}{[T^{-1} \sum_t x_t^2]^{1/2}}$$

Entonces bajo la hipótesis nula de  $\varphi(i, j) = 0$ ,  $\sqrt{T} \hat{\varphi}(i, j)$  se distribuye asintóticamente como una normal con media cero y varianza que se puede estimar:

$$V \equiv \frac{T^{-1} \sum_t x_t^2 x_{t-i}^2 x_{t-j}^2}{[T^{-1} \sum_t x_t^2]^3}$$

Esta prueba está construida con el objetivo de rechazar la no linealidad aditiva y si combinamos con el estadístico BDS elaborado por Brocks en el año 1996 para probar la independencia en los datos.

Si describimos este estadístico utiliza el concepto de dimensión de correlación, definimos la función de correlación entre dos puntos como:

$$C_n(k) = \lim_{T \rightarrow \infty} \frac{1}{T^2} \sum_{i < j} \theta(k - \|x_i - x_j\|)$$

Donde  $\theta$  es la función heavyside también conocida como función indicadora y  $\|\cdot\|$  la norma suprema.



Grassberger y Procaccia desarrollan un estadístico que se distribuye como una normal estándar asintóticamente:

$$J_{n,T}(k) = \sqrt{T} \frac{C_{n,T}(k) - C_{1,T}(k)^n}{\hat{\sigma}_{n,T}(k)}$$

Bajo la hipótesis nula

$$C_n(k) \rightarrow \prod_{i=0}^{n-1} \text{Prob}[\|x_{t+i} - x_{s+i}\| < k]$$

, cuando  $T \rightarrow \infty$ , el estadístico prueba la hipótesis nula de independencia aleatoria y sistemas distribuidos idénticamente, por lo cual si rechazamos la hipótesis nula entonces sería consistente con algún tipo de dependencia en los datos.

## 5.4. Apéndice 2

### 5.4.1. Prueba de Ljung Box Portmanteau

Tomando la literatura(Cesar Andrés Ojeda Echeverri,2012).Box and Pierce(1970) propone el estadístico Portmanteau:

$$Q^*(m) = T \sum_{l=1}^m \hat{\rho}_l^2$$

Para probar la hipótesis nula  $H_0: \rho_1 = \dots = \rho_m = 0$  contra la hipótesis alternativa  $H_a: \rho_i \neq 0$  para algún  $i \in \{1, \dots, m\}$ . Si  $r_t$  es una sucesión independiente e idénticamente distribuidas. Si  $Q^*(m)$  es una variable aleatoria chi-cuadrado con  $m$  grados de libertad. Luego de la modificación de Ljung and Box(1978) al estadístico se tiene:

$$Q^*(m) = T(T+2) \sum_{l=1}^m \frac{\hat{\rho}_l^2}{T-l}$$

Por lo tanto la regla de decisión es rechazar  $H_0$ , en caso  $Q(m) > \chi_\alpha^2$  donde  $\chi_\alpha^2$  denota el percentil  $100(1-\alpha)$  de una distribución chi-cuadrado con  $m$  grados de libertad

## 5.5. Apéndice 3

### 5.5.1. Prueba de normalidad de Jarque Bera

Tomando la literatura(Cesar Andres Ojeda Echeverri,2012).La prueba de Jarque Bera por sus propiedades óptimas de potencia asintótica(Gonzales y Nieto,2008) es una de las pruebas de normalidad mas utilizadas comparando el tercer y cuarto momento de la distribución empírica con los teóricos de la distribución normal. Entonces si  $\{r_t, \dots, r_t\}$  es la serie de tiempo con  $t$  observaciones. La estadística de prueba:

$$JB = \frac{\hat{S}^2(r)}{\frac{6}{T}} + \frac{(\hat{K}(r) - 3)^2}{\frac{24}{T}}$$

$T$ =número de observaciones de la muestra,  $\hat{S}(r) = \frac{1}{(T-1)\hat{\sigma}^3} \sum_{t=1}^T (r_t - \hat{\mu}_r)^3$  es el coeficiente de asimetría muestral,  $\hat{K}(r) = \frac{1}{(T-1)\hat{\sigma}^4} \sum_{t=1}^T (r_t - \hat{\mu}_r)^4$  es el coeficiente de curtosis muestral,  $\hat{\mu}_r = \frac{1}{T} \sum_{t=1}^T r_t$  es la media muestral y  $\hat{\sigma}_r^2 = \frac{T-1}{T} \sum_{t=1}^T (r_t - \hat{\mu}_r)^2$  la varianza muestral. JB es una variable aleatoria asintóticamente distribuida como una chi-cuadrado con dos grados de libertad.

## 5.6. Apéndice 4

### 5.6.1. Criterios de información

Tomando la literatura(Cesar Andres Ojeda Echeverri,2012).Los criterios de información tiene una estructura de complejidad y capacidad predictora es decir precision y exactitud por lo tanto miden la capacidad explicativa de un modelo y penalizan por su grado de complejidad.Davidson(2001) define los siguientes criterios de información:

Akaike Information criterion (AIC) :

$$-C_T - p$$

Schwarz Information criterion (SIC):

$$-C_T - \frac{1}{2}p \log T$$

Hannan Quinn criterion (HQC) :

$$-C_T - p \log \log T$$

p es el número de parámetros a estimar y T el tamaño muestral.